Medicine

UNIVERSITY OF Southampton

NHS National Institute for Health Research

PATIENT-REPORTED OUTCOME MEASURES FOR MONITORING PRIMARY CARE PATIENTS WITH DEPRESSION: PROMDEP RANDOMISED CONTROLLED TRIAL

PROTOCOL

Chief Investigator: Prof Tony Kendrick MD FRCGP FRCPsych (Hon)
Professor of Primary Care
Primary Care & Population Sciences
University of Southampton
Aldermoor Health Centre
Southampton SO16 5ST
Tel: 02380 591790
Email: ark1@soton.ac.uk

Sponsor: **University of Southampton**

ERGO Ref: **42615**

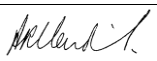
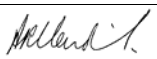
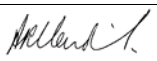
IRAS Ref: **250225**

Funder: **National Institute for Health Research Health Technology Assessment programme**
Ref: **HTA 17/42/02.**
Funding acknowledgement: The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

ISRCTN no: **17299295**

Dates: **01/11/2018 – 31/10/2021**

NHS Research
Ethics Committee ref: 18/WS/0144 approved by West of Scotland REC 5, 21st September 2018

Protocol Version	**Protocol date**	**Signature of CI**
1.1	12.09.2018	
1.2	24.10.2018	
1.3	07.11.2018	
1.4	04.01.2019	
1.5	04.05.2020	
1.6	**19.01.2021**	
1.7	**01.03.2021**	
1.8	**10.06.2021**	

Contents

MAIN STUDY CONTACTS

Chief Investigator: Professor Tony Kendrick
Primary Care and Population Sciences, Faculty of Medicine, University of Southampton
Tel: 02380 591790
Email: ark1@soton.ac.uk

Study Coordinator:
For general study and clinical queries e.g. participant queries, study supplies, data collection, please contact in the first instance:
Mrs Rachel Dewar-Haggart (Tel: 07880427831), Dr. Lien Bui (07920084086) or Miss Riya Tiwari (07789942927)
Email: promdep@soton.ac.uk.

SPONSOR

The University of Southampton is the research sponsor for this study. For further information regarding sponsorship conditions, please contact the Head of Research Governance at: Research Governance Office, University of Southampton, Room 4079, Building 37, Highfield Campus, Southampton SO17 1BJ
Tel: 023 8059 5058
Fax: 023 8059 5781
Email:rgoinfo@soton.ac.uk

CO-APPLICANTS

Prof Michael Moore
University of Southampton

Prof Geraldine Leydon
University of Southampton

Dr Beth Stuart
University of Southampton

Dr Adam Geraghty
University of Southampton

Dr Guiqing Lily Yao
University of Leicester

Prof Christopher Dowrick
University of Liverpool

Prof Glyn Lewis
University College London

Professor Gareth Griffiths
Southampton Clinical Trials Unit

Prof Carl May
London School of Hygiene and Tropical Medicine

Rachel Dewar-Haggart, Senior Research Assistant
University of Southampton

Samantha Williams, Senior Research Assistant
University of Southampton

Patient colleagues

Margaret Bell, PPI representative, Portsmouth

Bryan Palmer, PPI representative, Southampton

Research team, University of Southampton

Rachel Dewar-Haggart, Trial Coordinator
Lien Bui, Trial Coordinator
Riya Tiwari, Trial Coordinator

Sophie Johnson, Administrative Assistant

Dr Shihua Zhu, Health Economist

Louisa Little, Clinical Trials Manager, Southampton CTU

Research team, University of Liverpool

Emilia Trapasso, Trial Coordinator
Tasneem Patel, Trial Coordinator

Daniel Lawrence, Administrative Assistant

Research team, University College London

Jane Sungmin Hahn, Trial Coordinator
Mekeda Logan, Trial Coordinator

Tom Gant, Administrative Assistant.

FUNDER

This study is funded by an NIHR Health Technology Assessment project grant, ref: 17/42/02.

PROTOCOL INFORMATION

This protocol describes the PROMDEP RCT and provides information about procedures for entering participants. The protocol should not be used as a guide for the treatment of other non-study participants; every care was taken in its drafting, but corrections or amendments may be necessary. These will be circulated to investigators in the study, but sites entering participants for the first time are advised to contact Primary Care at the University of Southampton to confirm they have the most recent version.

LAY SUMMARY

We want to look at whether giving personal feedback to people being treated for depression might help them get better more quickly. One way of doing this is by using patient reported outcome measures (or 'PROMs') which involve patients filling out questionnaires to record their symptoms of depression and feeding back the questionnaire results to the health professionals looking after them, at follow-up appointments. Some benefit for patients from reduced depression has been shown to result from monitoring their progress with PROMs, at least in specialist psychological therapy and mental health settings, but this approach has not yet been researched in UK general practices. General practice is the setting in which most people with depression are treated in the UK, so it's important to test whether PROMs can be helpful in that setting.

In a previous study in general practices in southern England between 2014 and 2016, we found that it was feasible to allocate patients at random to either assessment with PROMs or usual care, so we could test the value of follow-up monitoring of depression with PROMs. The practices recruited patients with depression through sending letters to patients who had recently been seen for a new episode of depression, which proved an acceptable method to practices and patients. Patients were happy to complete our chosen questionnaires on symptoms, daily activities, quality of life, satisfaction with care, and use of GP and specialist services, and most attended follow-up as requested. Lower levels of depression symptoms at 12 weeks among patients in the PROM group suggested that completing PROMs at diagnosis and follow-up assessment may improve the outcome of depression treatment for patients, even though some GPs did not use the results to inform patient care.

The lessons learned from the previous feasibility study have been taken into account in the design of the larger definitive study described in this protocol. We will provide more engaging and instructive training materials for GPs using PROMs, indicating specific actions to be taken following assessment. We are using the PHQ-9, a PROM that is acceptable to GPs and identified as useful by them. Most GPs we talked to have used the PHQ-9 questionnaire, found it helpful where the diagnosis was uncertain, and feel they understood the implications of scores for their treatment decisions. We will also provide written feedback on PROM results for patients, listing possible treatment options. As patients do not always see the same GP at follow-up as they did at diagnosis, all the GPs in a practice will be trained to follow the same procedures in their appointments with people with depression, to make sure they can stick to what we are asking them to do in the study. We will identify and recruit practices with a track record of recruiting patients with depression, and ensure all the GPs are trained in the study procedures, including new GPs who join practices during the study. Rates of following up patients will be maximised through obtaining their permission to text, phone and email them to maintain better contact, more often.

This is an important study. If using PROMs is helpful in improving patient outcomes for depression then they are likely to be good value for money, given their low cost. The benefits could be considerable, given that depression is so common, can be so disabling, and costs the nation billions of pounds.

ABSTRACT

Design: Parallel group cluster randomised trial with 1:1 allocation to intervention and control.

Inclusion: Patients age 18+ years, with a new episode of depressive disorder/symptoms. If possible, patients will be recruited opportunistically at consultations for new episodes of depression. Records will also be searched for newly presenting patients, to reduce the risk of selection bias.

Exclusions: current depression treatment, comorbid dementia/psychosis/substance misuse/suicidal ideas.

Intervention: Administration of PHQ-9 soon after diagnosis, and at follow-up 10-35 days later. We will target GP reflective motivation and psychological capability with guidance on assessment and treatment, informed by NICE guidelines. GPs will be trained in interpreting scores, along with asking open-ended questions and exploring the patient's life context, and asked to take them into account in their treatment decisions. Patients will be given written feedback on their scores and suggested treatments to discuss with GPs. Control practice patients will not complete the PHQ-9. They will complete research outcome measures but not be given feedback on the results.

Baseline measures: depression on the Beck Depression Inventory BDI-II, sociodemographics, duration and past history of depression, GAD-7 for anxiety.

Outcomes: Primary outcome: BDI-II at 12 weeks. Secondary outcomes: BDI-II at 26 weeks; changes in drug treatments and referrals; social functioning (Work & Social Adjustment Scale) & quality of life (EQ-5D) at 12 & 26 weeks; service use over 26 weeks (modified Client Services Receipt Inventory) to calculate NHS costs; patient satisfaction at 26 weeks (Medical Informant Satisfaction Scale).

Sample size: Assuming baseline mean BDI-II 24.0; SD 10.0 (from feasibility RCT); follow-up mean of 14.0 at 12 weeks in intervention group, 17.0 in controls (difference 3.0 = effect size of 0.3 and MCID of 17.5% of control group score); mean 6 patients per practice; ICC 0.03; 5% significance; 90% power; correlation between baseline and follow-up of 0.5, needs 222 patients analysed per group. Cluster design effect 1.15; assuming 20% loss to follow-up gives $222x1.15x2/0.8 \times 1\text{-}0.5^2$=554 total, from 113 practices across three centres.

Randomisation: by CTU statistician with computerised sequence generation.

Blinding: of practitioners and patients is impossible given the nature of the intervention. Self-report outcome measures will prevent researcher rating bias.

Analysis: Differences at 12 and 26 weeks between intervention and controls in depression, social functioning and quality of life will be analysed using linear mixed models, adjusted for sociodemographics, baseline depression, anxiety, and clustering, including practice as a random effect. Patient satisfaction, quality of life (QALYs) and costs over 26 weeks will be compared between arms.

Qualitative process analysis: Interviews with 15-20 GP/NPs and 15-20 patients per arm to reflect on trial results and implementation issues, using Normalization Process Theory as a framework for the interview schedules and qualitative analyses. Practitioner/patient dyads to be interviewed as soon as possible after patient assessments at follow-up consultations, to explore recall of practitioner-patient discussion of scores and identify variations in the use of the PHQ-9.

BACKGROUND

The problem being addressed

England, like other countries, has seen big increases in antidepressants and psychotherapy for depression since the early 90s, yet the prevalence of depression has not reduced, but actually increased slightly. One of the main reasons is a lack of application of evidence-based treatments to those who would benefit– the 'quality gap' (1).

NICE guidelines recommend different treatments for more severe depression than for less severe depression (2). However, GPs, who treat more than 80% of cases in primary care, are often inaccurate in their global clinical assessments of depression severity, and so treatment is not targeted to patients most likely to benefit (3). Some patients receive treatment they don't need (medicalising self-limiting illness and exposing them to side effects) and others don't get treatment they do need, significantly contributing to the 'quality gap'. A systematic review concluded there are many false diagnoses as well as missed cases, which could be improved by reassessment of individuals who might have depression (4). As a result, NICE (2) recommends practitioners consider using depression symptom questionnaires as validated measures of severity at diagnosis and follow up, to inform and evaluate treatment.

Questionnaire use was incentivised in the GP contract from 2006-2013. However, since those payments stopped, most GPs prefer not to use them, saying they intrude in consultations and undermine their autonomy. Some doubt their validity, preferring to use their own judgement to assess severity and response to treatment (5).

A Cochrane review of using PROMs in treating common mental health disorders (CMHDs) including depression found some evidence of benefit in psychotherapy and specialist mental health settings, but the research was generally of low quality and hardly any research had been done in primary care (6). More research is therefore required, particularly in UK primary care, where most patients are treated if they are treated at all.

Importance of the research

Depression is common and costly. The 1-week prevalence among adults in the UK is 11.1%, including 3.3% major depression and 7.8% mixed depression and anxiety (7). It can lead to chronic disability, poor quality of life, suicide, and high service use and costs. The King's Fund estimated that 1.45 million people will have depression in England by 2026, and annual costs for care, social services and lost employment will be £12.2 billion (8). If using severity measures improves the targeting of treatment and outcomes for depression even to a modest extent, they are likely to be cost-effective given their low cost, and the benefits at a population level would be considerable in public health terms, given the high costs of depression.

Depression symptom questionnaires are an example of patient-reported outcome measures (PROMs), which have been promoted to increase patient involvement in their care (9), and research shows that patients value the use of questionnaire severity measures to confirm their diagnosis and monitor their progress (5).

Observational research suggests depression questionnaires can also improve the process of care for patients. Following NICE guidance, from 2006-2013 the GP contract Quality and Outcomes Framework (QOF) paid GPs to use symptom questionnaires to assess depression severity at diagnosis of a new episode. Questionnaire assessments at follow-up were also rewarded in the QOF from 2009-2013. Our previous observational research conducted in the year following the introduction of the QOF incentivisation of questionnaire use found that patients valued the use of them to confirm their diagnosis and monitor their progress, and some GPs also valued them for monitoring patients(5). Importantly, treatment was better targeted. The likelihood of antidepressant treatment or referral to psychology was

found to be significantly associated with higher questionnaire scores at diagnosis (10), and at follow-up decisions to change treatment were significantly associated with lack of improvement in scores(11).

The use of questionnaires was disliked by some GPs however, saying they intruded in consultations and undermined their autonomy. Some doubted their validity, preferring to use their own judgement to assess severity and response to treatment (5). In response to criticisms NICE commissioned a review (12) which concluded the evidence was not strong enough to require their use in QOF depression indicators. Currently the QOF rewards reviews 10–35 days after diagnosis but questionnaires are optional and not required to receive payments.

However recent time series analysis of GP prescribing data shows the QOF depression indicators were associated with a subsequent reduction in antidepressant prescribing for first-ever episodes of depression, in line with NICE guidance not to give antidepressants for mild depression(13), so rewarding questionnaire use should be seriously reconsidered.

Routine outcome monitoring with questionnaire measures of depression severity takes place in the NHS Improving Access to Psychological Therapy (IAPT) psychological treatment services, and has been shown to improve the efficiency of care in that setting(14). However, only 15% of CMHD patients are treated by the IAPT programme (15), so more research into the potential benefits of routine outcome monitoring with depression symptom severity measures is required in UK primary care, where the large majority of patients are treated if they are treated at all.

Previous research

Systematic reviews of PROMs for depression have found some evidence of benefit for patients treated in mental health (16) and psychological therapy settings (17), but a recent Cochrane review found little research had been done in primary care (6).

We carried out a feasibility RCT of PROMs for depression in UK primary care (18). We tested individual patient and cluster randomisation, in 9 practices, of 47 adults with new episodes: 22 intervention, 25 control. Three PROMs were administered following diagnosis and again 10-35 days later: the Patient Health Questionnaire PHQ-9(19), Distress Thermometer analogue scale (20), and PSYCHLOPS problem profile (21). Feedback of scores to patients was left to the practitioners. Mean BDI-II score at 12 weeks was lower among intervention group patients than controls by 5.8 points (95% CI-11.1, -0.5), adjusted for baseline differences and practice (18). Social functioning scores were not significantly different. At 26 weeks there were no significant differences in symptoms, social functioning, quality of life, or costs, but mean satisfaction score was lower among intervention patients by 22.0 points (-40.7,-3.29). Qualitative interviews suggested this was because patients were disappointed when their GPs did not use PROM scores to inform treatment. Some GPs were not convinced the PROMs were useful and wanted more guidance on treatment actions in response to the scores (18).

We concluded PROMs might improve depression outcomes, even if they do not always inform management, in line with the findings of a similar trial using the PHQ-9 in the USA (22). Patients can feel more involved in their care and more motivated to adhere to treatment and follow-up (23, 24). Primary care patients in Sweden who were monitored with the Montgomery-Asberg rating scale were more likely to adhere to antidepressants, although there was no improvement in outcome in that study (25), so findings are variable and more research is needed in UK primary care.

AIM

The aim is to answer the following research question: What is the effectiveness and cost-effectiveness of assessing primary care patients with depression or low mood soon after diagnosis and again at follow-up 10-35 days later, using the PHQ-9 questionnaire combined with patient and practitioner feedback and guidance on treatment?

OBJECTIVES

The objectives are:

1. To carry out a cluster randomised parallel group controlled trial that will compare (i) getting patients to complete the PHQ-9, for use as a patient reported outcome measure (PROM) in their consultations with GPs or Nurse Practitioners (NPs) treating them for depression, with (ii) usual practitioner care, uninformed by PHQ-9 scores.

2. To motivate and train participating practitioners to reflect on the best use of the PHQ-9, improving their capability to interpret symptom scores, taking into account patients' responses to open-ended global enquiries, their level of functioning, past history, and social context including life events and difficulties.

3. To provide patients in the intervention arm with written feedback on their PHQ-9 scores, including a 'traffic light' indication of the level of severity of their depression, a 100 manikin representation of the proportion of people in the population with that level of depression, and a brief list of evidence-based treatments relevant to the level of severity, which they will be asked to discuss with their GP/NP.

4. To follow up participants for 26 weeks, with research assessments at 12 and 26 weeks.

5. To determine the primary outcome of depressive symptoms on the Beck Depression Inventory, 2nd edition (BDI-II), at 12 weeks follow-up.

6. To examine secondary outcomes including depressive symptoms on the BDI-II at 26 weeks, and social functioning, quality of life, and changes in drug treatment and referrals, at both 12 and 26 weeks follow-up.

7. To measure service use and costs over the 26-week follow-up period and perform cost-effectiveness and cost-utility analyses based upon the results of the trial.

8. To carry out a qualitative process analysis to explore participants' reflections on the conduct of the trial, and the potential for implementing the use of PROMs in practice. Interviews with 15-20 practitioners & 15-20 patients will be carried out, using Normalization Process Theory (26) as a framework for the interview schedules and qualitative analyses. Practitioner/patient dyads will be interviewed as soon as possible after patient assessments at follow-up consultations, to explore recall of interactions with their GP/NPs and identify variations.

METHODS

A randomised parallel group controlled superiority trial that will compare (i) getting patients to complete the PHQ-9 questionnaire as a patient reported outcome measure (PROM) for use in practitioner consultations for depression, with (ii) usual care, uninformed by PHQ-9 questionnaire scores.

Intervention

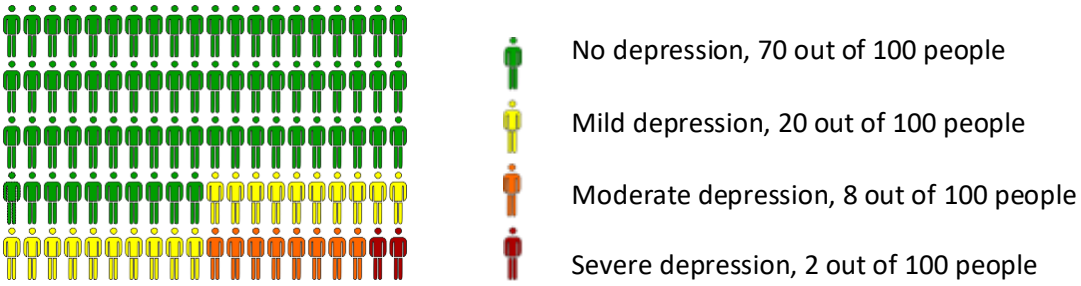
The intervention consists of getting patients to complete the Patient Health Questionnaire (PHQ-9) which measures depression symptoms, for use as a patient reported outcome measure (PROM) in their consultations with GPs (or Nurse Practitioners, NPs) treating them for depression. The PHQ-9 will be completed by participating patients as soon as possible after diagnosis, and then again at a follow-up consultation 10-35 days after that (this follow-up time period has been chosen as it is the interval laid down for financially incentivised follow-up assessments in the GP contract quality outcomes framework QOF). Patients will be given feedback on the meaning of their symptom score and possible treatment options to discuss with the practitioner. Practitioners will be trained in the interpretation of symptom scores in the context of the patient's life situation, and in further assessment to inform their treatment decisions.

The PHQ-9 is a nine question self-report measure of depression symptoms that takes approximately three minutes to complete (19). It asks about the American Psychiatric Association's Diagnostic and Statistical Manual (DSM) nine diagnostic symptoms of major depressive disorder, and scores on each symptom range from 0 (not at all) to 3 (nearly every day). Total scores are categorised into minimal or no depression (0-4), mild (5-9), moderate (10-14), and severe (15-27). It was developed and originally validated against diagnostic interviews in the USA and can be downloaded free of charge from: www.depressionprimarycare.org/clinicians/toolkits/materials/forms/phq9/questionnaire/
Pfizer owns the copyright but does not charge for its use in clinical practice or research (27).

Patient feedback on the meaning of their PHQ-9 score

We will provide patients in the intervention arm with written feedback on their PHQ-9 scores, including a 100 manikin representation of the proportion of people in the population with that level of depression, a 'traffic light' indication of the level of severity of their depression, and a brief indication of possible evidence-based treatments relevant to the level of severity, which they will be asked to discuss with their GP/NP. PHQ-9 sum scores range from 0 to 27, with scores of 0-4, 5-9, 10-14 and 15+ representing probable minimal or no depression, mild, moderate and severe depression symptom levels, respectively. These will be fed back to patients in four probability categories: green, yellow, orange and red respectively. An example of feedback to a patient scoring 14 on the PHQ-9 is shown below.

Distribution of depressive symptom severity in the general population



You fall into this category: moderate depression.

Orange: you probably have moderate depression

If you have moderate depression, you may benefit from referral for psychological therapy (talking treatment) or from antidepressant treatment from your GP or nurse practitioner, or both. Please ask your GP about these treatment options when you go back for your next appointment at the practice.

This approach proved successful in Löwe *et al*'s DEPSCREEN-INFO study of providing written feedback after screening patients with cardiological problems for depression (28). (Löwe is an international advisor on our proposal). Six months after screening, the patient-feedback group showed significantly greater improvements in depression severity and was twice as likely to seek information about depression compared with the control group.

GP/NP training and guidance in use of PHQ-9

The PHQ-9 meets minimum standards for a PROM (29) in terms of established validity in UK primary care (30) and sensitivity to change in response to treatment, at least at the group level (31). It was recommended for use by the PROM Research Group at the Oxford Department of Public Health as having evidence to support its use and being broadly acceptable (32). Many GPs in UK practice are familiar with the PHQ-9, as it was the most frequently used PROM in practices in the period (2006-13) during which the use of depression symptom questionnaires was incentivised by the QOF (10). However, some GPs doubt the validity of the PHQ-9, preferring to use their own judgement to assess severity and response to treatment (5).

The severity categories of the PHQ-9 can be criticised, as a score of 5-9 ('mild depression') is found in around 20% of the population but is not usually associated with significantly impaired functioning, so labelling people scoring 5-9 as having depression may be counter- productive, in that treatment is not usually indicated and the label itself may make patients feel worse about themselves. The cut-off of a score of 10 for moderate depression, as the threshold for offering treatment, has also been questioned, and studies have suggested a score of 12 may be a more valid threshold (10). The PHQ-9 tends to put more people in the 'moderate depression' category than other PROMs such as the Hospital Anxiety and Depression Scale for example (10, 33).

In any case, an initial high PHQ-9 score by itself, whether above 10 or above 12, does not suffice to indicate the need for antidepressant drug treatment or referral for psychological therapy, because patients vary a lot in their propensity to acknowledge symptoms when asked. Recent qualitative research led by co-applicants Dowrick and Lewis suggests that the PHQ-9 is not exhaustive in its list of symptoms and not all patients find it straightforward to complete, so it may miss symptoms that are meaningful to patients (e.g. changes in libido, social withdrawal, interpersonal difficulties), and underestimate their intensity (34). Some patients consciously under-report their symptoms to try to reduce them through positive affirmations, while others over-report them to emphasise they want help. Consequently, as many as half of patients rated with the PHQ-9 may have a mismatch with how they describe their overall condition at baseline, as well as progress over time, when asked global, open-ended questions (34).

Symptom scores are therefore quite individual and the baseline level has to be interpreted in light of the impact of the symptoms on the person's functioning at home, at work, and in relationships. The recent qualitative research suggests that the PHQ-9 should not be used as a standalone tool but should preferably be used in conjunction with an open-ended enquiry such as 'how are you feeling in yourself?' (34), as a better measure of the person's unique ongoing

experience of their depression. Additionally, within-person changes in individuals' PHQ-9 scores between the first and second consultation are only limited indicators of whether patients are improving, not improving, or getting worse, and therefore need to be supplemented with a global enquiry such as 'how are you feeling in comparison to when I last saw you?' along with an update on their life circumstances (34). Where there are mismatches between changes in patients' scores and their global ratings of change, practitioners need to take particular care when interpreting the results of the PHQ-9.

Our feasibility study suggested that GPs' discussion of the PHQ-9 scores with patients, and use of them to inform treatment, were suboptimal, affecting both their own perception of the measure, and patients' satisfaction with the care they received (18). To change practitioner behaviour in the proposed trial, we will implement two hours of structured training. By triangulating our qualitative feasibility findings with behavioural theory (35), we determined the need for the training to focus primarily on GP's reflective motivation (e.g. beliefs about the usefulness of PROMS) and psychological capability (e.g. knowledge and understanding to apply PROMS effectively). These constructs are drawn from the 'COM-B' system of behaviour (referring to Capability, Opportunity, Motivation and Behaviour) (35). The COM-B system is used widely in behaviour change research, and focuses on necessary antecedents for voluntary behaviour to occur.

Participating GPs will therefore be given two hours of training either face to face on their practice premises, or on-line, including written material beforehand, a PowerPoint presentation, case vignettes, and questions for them to answer to show they have understood the training. The training will focus on evidence that patients do value using PROMs and can benefit from being more involved in their own care even if the scores don't alter treatment, to get GPs/NPs to reflect on the value of the use of the measure. We will address GP/NP concerns around the validity of the PHQ-9 by acknowledging individual differences in patient response set, and advising them to combine more global open-ended questions with the questionnaire measure.

Setting

The study will be carried out in primary care, recruiting general practices around three sites: the University of Southampton, University of Liverpool, and University College London.

Design

The study design, informed by the successful feasibility trial, is a parallel group cluster randomised trial, with patients clustered by participating practices, and 1:1 allocation of practices to intervention and control groups. We have chosen a cluster-randomised design, as randomising patients individually within practices risks contamination between study arms. GP/NPs taught to use a symptom questionnaire with intervention arm patients may then use similar questions in a systematic way with control patients. In addition, patients do not always see the same GP/NP at diagnosis and follow-up, so all practitioners in a practice need to follow the same protocol, to optimise adherence to intervention or control arm procedures.

Randomisation

Randomisation of practices to intervention or control group will be carried out by a statistician from the Southampton clinical trials unit, so that the study statistician Stuart can remain blind to allocation. Randomisation will be by computerised sequence generation, and minimisation using three factors to avoid imbalance between the two arms: practice size (large vs small), location (urban/suburban vs rural), and centre (Southampton vs Liverpool vs UCL).

Blinding

Blinding of both patients and practitioners in the intervention arm is impossible given the nature of the intervention. Self-report outcome measures will therefore be used, to prevent observer rating bias by research team members aware of the patient's assigned trial arm.

Consent

For the baseline assessments, if the patient gives verbal and then subsequent written consent to take part (by post or online), the researcher will conduct a baseline assessment. If online consent is given, this will be by completing an electronic copy of the consent form on Microsoft Forms.

For the patient and health professional qualitative interviews, participants will give verbal consent prior to the interviews. Consent will be audio recorded prior to the interviews, and saved as a separate file to the interview to ensure anonymity. Subsequent written consent will be obtained by completing an electronic copy of the consent form on Microsoft Forms.

Target population

The target population is patients aged 18 or more years, diagnosed by a GP or NP with a new episode of depression disorder or depressive symptoms. A new episode means no diagnosis or treatment within the previous three months.

Inclusion criteria

The main inclusion criteria will be adult patients seen in the practice within the last two weeks and assigned Read codes by GPs or NPs for new presentations with diagnoses or symptoms of depression. There will be no upper age limit, and no exclusion of patients with coexisting physical health problems.

Exclusion criteria

Patients will be excluded if they are already being treated for depression, or if they have comorbid dementia, psychosis, or substance misuse (as a main problem). Patients will also be excluded if they have significant suicidal thoughts requiring possible urgent referral to specialist mental health care (see below).

Patient recruitment and consent

Method 1

Where possible, patients who are seen with a new episode of depressive symptoms or disorder will be recruited opportunistically during consultations by participating GPs and NPs in both arms of the study. Patients identified through this method will be given the information sheet by hand, together with a reply slip and a Freepost envelope, and asked to contact the study team if they wish to take part.

Method 2

Method 1 may be subject to selection bias by the GP/NP however, and so patients presenting with a new episode of depressive symptoms or disorder will also be identified through weekly searches of practice medical records databases, to identify patients who were not selected by the GP/NP. In the feasibility trial both methods were used and 79% of patients were recruited in consultations opportunistically, and 21% through the weekly database searches, but this varied by practice and some practices recruited the majority of patients through the weekly searches.

Our experience gained recruiting people with depression for previous studies has shown that there are around 120 Read codes used by GP/NPs including both diagnostic codes (e.g. *major depressive disorder*) and symptom codes (e.g. *low mood*). Practices will use the full list for searching their databases weekly. Patients identified through this method will be mailed an information sheet about the study by the practice and asked to contact the study team if they wish to take part, or to decline, using a reply slip and a Freepost envelope. If they do not respond the research team will have no knowledge of them, maintaining patient confidentiality.

Telephone screening prior to recruitment

If patients do respond positively to either approach, a member of the research team will then contact them, screen them by telephone for any exclusion criteria, and arrange to see them face to face or remotely, using Skype or MS Teams, telephone, or another communication platform for the baseline visit if they are eligible.

Baseline visit

The baseline visit will be offered either at their general practice premises, at their home, or remotely, depending on patient preference, or necessity during the Coronavirus crisis. They will attempt to meet the patient within a week of receiving their reply slip indicating their interest in participating, in order to see them within 2-3 weeks of their initial presentation to the GP/NP. At the initial contact the researcher will go over the patient information sheet again, seek verbal consent, and carry out the baseline research assessments. The participant will be asked to confirm their consent in writing afterwards.

Procedures in intervention and control arms

Intervention arm

The PHQ-9 will be administered as a PROM at baseline by a member of the research team, and at the follow-up consultation by the practitioner treating the patient. (In practice, as opposed to the trial situation, the GP/NPs themselves would ask the patients to complete the PROM, either in the first consultation for depression, or between consultations, but the trial situation is different, as patients have to consent to take part after being given sufficient time to consider this, so the GP/NP cannot give a PHQ-9 questionnaire at the first consultation. We will take this difference into account when modelling the costs of the intervention).

After baseline assessment, the researcher will ask the patient to complete the full PHQ-9 questionnaire, either on paper or online. The patient will then be given an information sheet by the researcher tailored for their particular PHQ-9 score, which will include the 100-manikin representation and traffic light indication of severity, and suggestions about possible treatment, as described above. The patient will then be asked to arrange an appointment with their GP/NP, either remotely or in person as soon as possible and to take with them their completed PHQ-9 questionnaire plus written feedback, in order that they can discuss the score and the treatment suggestions with the GP/NP. To ensure the PHQ-9 result and patient feedback reaches the practice we will email it using a secure NHS email account (we have done this successfully in a previous trial).

Participating GPs/NPs will be asked to take the PHQ-9 scores and patient advice into account when deciding about treatment at their next consultation with the patient, following the treatment guidance given during training, taking the patient's response to a global open-ended inquiry into account, together with their level of functioning, social context and past history.

The GP or PN will also be asked to provide the patient with a fresh PHQ-9 at that second consultation for the patient to take away and complete immediately prior to a third, follow-up consultation 10-35 days later. At that third

consultation the GP/NP will be asked to go through the follow-up PHQ-9 with the patient and take the change in score between consultations into account when deciding about possible changes to treatment.

Control arm

In control arm practices, patients will not complete the PHQ-9. They will be seen by the research team either remotely or in person as soon as possible after their first consultation with depressive symptoms, and asked to complete baseline research outcome measures, but will not be given feedback on the results of those. They will be asked to arrange a follow-up appointment with the GP/NP, either remotely or in person to match what happens in the intervention group, but the GP/NP treating them will not receive training, and will be asked to provide their usual care.

Timing of starting treatment

Practitioners in both the intervention and control groups will be advised that best practice in treating depression is not to start treatment at the consultation at which symptoms of a new episode are presented by the patient, unless they think it is absolutely indicated in their clinical judgment. This is because a significant proportion of patients will improve without treatment within 2-3 weeks, having had their problems acknowledged and having received general advice about the nature and course of depression. We are interested in this study with the use of the PHQ-9 in deciding on initial treatment, as well as follow-up monitoring, so we prefer treatment is not started before the baseline assessment in both groups, and before the first PHQ-9 questionnaire is administered by the researcher in the intervention group. In the feasibility study this was carried out on average 10 days (range 1-38 days) from receiving the patient's reply slip, and it should be possible to complete baseline assessment within two weeks of the patient's first presentation in most cases.

It is possible however that patients recruited either opportunistically or via the weekly searches will have been started on treatment at the consultation when they first presented with a new episode, if treatment cannot be postponed in the judgement of the treating practitioner. We will record whether treatment has already started at the baseline assessment.

Assessment measures at baseline and follow-up

Patients will be recruited over a 24-month period and followed up for 26 weeks each, with assessments at baseline, 12 and 26 weeks. Follow-up assessments will take place either at their general practice, at their home, or remotely if they prefer, or if necessary, due to the Coronavirus crisis.

At the baseline visit, the following measures will be administered:

- the Beck Depression Inventory second edition BDI-II (36) for current level of depression
- a bespoke questionnaire on sociodemographic details (age, gender, ethnicity, education, employment, and cohabitation status)
- a bespoke questionnaire on the duration of the current episode of mood disturbance, and any past history of depression
- the 7-item generalised anxiety disorder (GAD-7) questionnaire for anxiety symptoms (37)
- the Work & Social Adjustment Scale (38) for social functioning
- the EuroQol 5-item 5-level (EQ-5D) questionnaire for quality of life (39)
- bespoke questionnaires on consultations, drug treatments and referrals for depression over six months (to calculate NHS costs)

At 12 weeks follow-up the following measures will be administered:

- the BDI-II to measure changes in depressive symptoms (primary outcome)
- the Work & Social Adjustment Scale to measure changes in social functioning
- the EQ-5D to measure changes in quality of life

At 26 weeks follow-up the following measures will be administered:

- the BDI-II to measure changes in depressive symptoms (secondary outcome)
- the Work & Social Adjustment Scale to measure changes in social functioning
- the EQ-5D to measure changes in quality of life
- bespoke questionnaires on consultations, drug treatments and referrals for depression over six months (to calculate NHS costs)
- a modified version of the Medical Informant Satisfaction Scale MISS (41) to measure patient satisfaction over the whole 26 weeks

Data collection will be through face-to-face meetings or remotely, on-line, or by post, but brief telephone follow-up will be offered to obtain at least the primary outcome (BDI-II score) and quality of life (EQ-5D) for the study's primary clinical and economic outcomes, if the researcher is unable to arrange to assess the patient fully on-line, by post, or face to face.

Participants will receive a £10 high street shopping voucher at both the 12 and 26 week follow-ups, to thank them for their participation in the study.

The Beck Depression Inventory, second edition BDI-II is a 21 item self-report instrument that uses DSM-IV criteria (36). It has been established as a valid and reliable instrument for depression screening in the general population (36, 42) and is widely used in depression trials. It takes approximately five minutes to complete. Each item is scored from 0-3 and a total score of 0-13 is considered minimal range, 14-19 is mild, 20-28 is moderate and 29-63 is severe.

The GAD-7 score is a 7-item measure of anxiety symptoms (37). The total score is calculated by assigning scores of 0, 1, 2, and 3, to the response categories of 'not at all', 'several days', 'more than half the days', and 'nearly every day', respectively, and adding together the scores for the seven questions. Scores of 5, 10, and 15 are taken as the cut-off points for mild, moderate and severe anxiety, respectively. Using the threshold score of 10, the GAD-7 has a sensitivity of 89% and a specificity of 82% for generalised anxiety disorder (37).

The Work and Social Adjustment Scale assesses problems in functioning with work, home management, social leisure activities, private leisure activities, and family & relationships, all on 0 to 8 scales (38). It has been shown to be a sensitive, reliable and valid measure of impaired functioning and is used routinely in IAPT psychological therapy settings as well as in research studies in a variety of settings.

The EuroQol-5D (EQ-5D)-5L measure of health-related quality of life (39) is the measure favoured by NICE in determining cost-effectiveness when developing its clinical guidelines. The EQ-5D includes five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression, each scored on five levels. Health states are converted into a single summary index by applying weights to each level in each dimension derived from the valuation of EQ-5D health states in adult general population samples (43). The EQ-5D measure of patient utility will be used to determine changes in quality adjusted life years (QALYs) for the health economics evaluation.

Costs will be calculated from responses to the Client Service Receipt Inventory CSRI (40), modified specifically for the study. A review of participating patients' digital medical records will also be carried out by practice staff after the 26

week follow-up, to augment questionnaire measurement of health and social service resource use using the modified CSRI.

Patient satisfaction will be assessed using the 29-item 'Medical Interview Satisfaction Scale' MISS-29 which was developed in the USA to assess patient satisfaction with individual doctor-patient consultations, and has been shown to be valid and reliable in UK primary care (41). We will adapt it to rate patient satisfaction at the 26-week follow-up, asking patients to look back over their consultations with GPs/NPs over the whole 26 week period.

Assessing suicidal ideation

Patients who disclose information during an interview (face-to-face, over the telephone, or remotely) to the fieldworker indicating that they have attempted suicide or that they have been thinking of ways to commit suicide will be considered to have suicidal ideation. Their GP/NP will be informed immediately by telephone and asked to review the patient, with a view to a possible urgent referral to mental health services.

Sample size

We need a sample large enough to detect a difference between arms at follow-up of the minimal clinically important score (MCID) on the primary outcome, the Beck Depression Inventory 2nd edition (BDI-II).

Button et al (44) used data collected from three randomised controlled trials (n = 1039) for the management of depression, and compared improvement on a 'global rating of change' question with changes in BDI-II scores. They used general linear modelling to explore baseline dependency, assessing whether MCID is best measured in absolute terms (i.e. difference) or as percentage reduction in scores from baseline (i.e. ratio). The modelling indicated that MCID is best measured on a ratio scale as a percentage reduction of score, and an MCID of a 17.5% reduction from baseline was identified from receiver operator characteristics analyses as the optimal threshold above which individuals reported feeling 'better' (44).

In the PROMDEP feasibility trial we found the mean BDI-II score at baseline was 24.0 and the standard deviation (SD) was 10.0 (18). At 12 weeks follow-up, based on the results of the feasibility study, we anticipate a mean of 14.0 in the intervention group, and 17.0 in the control group. This gives a mean difference of 3.0 on the BDI-II, which is an effect size of 0.3 SDs, and in keeping with the findings of Knaup *et al*'s systematic review for the expected effects of combined practitioner and patient feedback of PROMs (16). The difference of 3.0 points is 17.6% of the control group's score of 17.0 at 12 weeks, and therefore just above the MCID for the BDI-II (44). The anticipated potential benefit would therefore be small, but clinically significant.

We will aim to recruit a mean of six patients per practice. We assume an intra-cluster correlation coefficient (ICC) of 0.03 (from the feasibility study). At the level of 5% significance, to have 90% power to detect a difference between 14.0 and 17.0 on the BDI-II we need 235 patients analysed per group. Given a cluster size of six, the cluster design effect will be 1.15, meaning we need 270 per group. We assume a 20% loss to follow-up at 12 weeks so the total sample size needed will be 270 x 2/0.8 so our original sample size was a total of 676 patients recruited, from 113 practices, across the three recruitment centres (around Southampton, UCL, and Liverpool).

We have however found a correlation between baseline and follow-up values for the primary outcome (the BDI-II score at 12 weeks) of p=0.6 (95% CI 0.5 to 0.7). Assuming conservatively that this correlation p remains 0.5 or greater until the end of follow-up, the necessary target sample size to give 90% power will be reduced by a deflation factor of $1\text{-}p^2$, i.e. $1\text{-}0.5^2$, which means we need 222 patients analysed per group. With the cluster design effect of 1.15; assuming 20% loss to follow-up gives 222x1.15x2/0.8 x (1-0.25) =554 total. (Revised 10th June 2021)

Data analysis

Primary and secondary outcomes

The primary outcome, differences at 12 weeks between intervention and controls in depression as measured by the BDI-II, will be analysed using a linear mixed model, adjusting for socio-demographics, baseline depression, anxiety, and clustering, including practice as a random effect. The model will use all the observed data and makes the assumption that missing BDI-II scores are missing completely at random.

Analysis of secondary outcomes, BDI-II at 26 weeks, social functioning, patient satisfaction and quality of life score, will also be conducted using linear regression for continuous outcomes and logistic regression for dichotomous outcomes, again adjusting for socio-demographics, baseline depression, anxiety, and clustering, including practice as a random effect.

We will examine the structure and pattern of missing data and, if appropriate, will present a sensitivity analysis based on data imputed using a multiple imputation model. Data will be analysed on an intention-to-treat basis.

No interim analyses are planned. Full details of the analyses to be undertaken will be set out in a Statistical Analysis Plan to be approved by the Trial Steering Committee and Independent Data Monitoring Committee.

Quantitative process analysis

We will examine the effects of practitioner and patient engagement with the PROM intervention on patient outcomes, adjusting for baseline patient characteristics (severity of depression, past history of depression, past history of treatment, and sociodemographic factors). We will employ multi-level modelling to investigate how these factors relate to outcomes in mediation analyses.

Qualitative process analysis

Process evaluation is an important tool for understanding both the dynamics and the outcomes of clinical trials, and Normalization Process Theory NPT (26) is a conceptual toolkit developed for this purpose (45). NPT focuses on understanding the mechanisms that promote, and the factors that inhibit, sense-making, participation, action and monitoring by participants in implementation processes.

The objectives of the process evaluation in the trial are to identify, characterise, and explain the perspectives of patient and practitioner participants on the conduct of the trial, and to construct a taxonomy of factors affecting both the conduct of the trial and the potential for normalisation of the use of PROMs in everyday practice, outside of the trial situation. The analysis will enable the construction of an implementation framework of barriers and facilitators (patient and health system factors) that need to be taken into account in the use of PROMs in primary care practice.

Interviews will be carried out with 15-20 practitioners and 15-20 patients in each arm (total 30-40 of each) to explore their reflections on the conduct of the trial, and the potential for implementing the use of PROMs in practice, using NPT as a framework for the initial interview schedules and qualitative analyses. For the patient and health professional qualitative interviews, participants will give verbal consent prior to the interviews. Consent will be audio recorded prior to the interviews, and saved as a separate file to the interview to ensure anonymity. Subsequent written consent will be obtained by completing an electronic copy of the consent form on Microsoft Forms.
Participants will receive a £10 high street shopping voucher for taking part in the interview. Practitioner/patient dyads will be interviewed as soon as possible after patient assessments at follow-up consultations, to explore patient and practitioner recall of interactions within the consultation, and to identify variations in the use of PROMs and in usual

practitioner care. Interviews will be transcribed and emerging themes identified through inductive analysis, using the constant comparative method (46). We will draw on insights from the wide range of studies that have employed NPT, giving a basic structure to the topic guide to be written in advance of the interviews. However, we will also work prospectively and inductively to ensure that we identify, characterise and understand (i) disconfirming evidence and deviant cases, and (ii) processes that are not accounted for within NPT.

Health economic evaluation

A health economic evaluation will be undertaken from an NHS and PSS perspective with a sensitivity analysis from a societal perspective. The outcome will be expressed as incremental cost per point improvement in the BDI-II clinical outcome, and incremental cost per quality adjusted life year (QALY) gained (cost utility analysis). All items will be costed using appropriate data (e.g. PSSRU NHS and social care reference costs (47), with informal care costed at minimum wage level. The primary analysis will be at 26 weeks. Personal costs will include patient and carer time off work, personal expenses, and travel.

A generalised linear mix model will be used to estimate the differences in costs and QALYs (using the EQ-5D to calculate patient utilities), adjusting for baseline characteristics including depression history, quality of life, and sociodemographic factors. Where appropriate we will estimate incremental cost-effectiveness ratios (ICERs). We will estimate mean values and 95% percentiles using non-parametric bootstrapping, and use these to produce cost-effectiveness acceptability curves (CEACs). Major assumptions in the costing and QALYs analysis will be tested through sensitivity analyses. Modelling of the likely benefit if any of using PROMs in practice will include making assumptions about the extra time which would have to be taken for GPs/NPs to administer the initial PROM (rather than the researcher) in the non-trial situation, together with any payments that might have to be made to the practice, e.g. through the QOF, to incentivise the use of the PROM.

ETHICAL CONSIDERATIONS

Patients with mental health problems like depression may be more sensitive than others to the demands of participation in research, but the effects of the problems in a primary care population are not so severe as to interfere with patients' capacity to understand the information provided or to give informed consent, provided patients with suicidal ideas, psychotic symptoms, and dementia are excluded. We have ensured the study aims are relevant to patients and the public through PPI input to the design, and their involvement will continue throughout to ensure that participation is voluntary, that easily understood patient information is provided, and fully informed consent obtained. Confidentiality and freedom to drop out at any time (see below) will be ensured.

In obtaining and documenting informed consent verbally, followed up in writing, the researcher will comply with applicable regulatory requirements and adhere to the principles of Good Clinical Practice (GCP). Discussion of objectives, risks and inconveniences of the study and the conditions under which it is to be conducted will be provided to the participant by appropriately delegated staff with knowledge in obtaining informed consent and with reference to the patient information leaflet. This information will emphasise that participation in the trial is voluntary and that the participant may withdraw from the trial at any time and for any reason. The participant will be given the opportunity to ask any questions that may arise and provided with the opportunity to discuss the study with family members, friend or an independent healthcare professional outside of the research team and time to consider the information prior to agreeing to participate.

Independent peer review through the NIHR HTA panel has ensured scientific quality and rigor. Ethics Committee and HRA approvals will be obtained prior to commencement of work with patients and health professionals, and any subsequent issues will be addressed with the REC or HRA offices as necessary.

PATIENT AND PUBLIC INVOLVEMENT

We recruited Bryan Palmer, convener of Southampton Depression Alliance, and another service user Margaret Bell, to join the feasibility study group. They made suggestions for the protocol, including having no upper age range but excluding people with dementia, and including people with less good English, providing interpreters if necessary. They read and commented on participant information sheets and consent forms to ensure ease of understanding and readability. Margaret read through some participant and GP interviews and made valuable comments. Both contributed ideas for this proposed definitive trial, including giving patients greater understanding of the PROMs and responsibility for feeding back to GP/NPs, to involve them more in their own care. Both have read and commented on this application, in particular checking the plain English summary. We also showed depression PROMs to Bryan's MIND/Depression Alliance self-help group, and the BDI was actually the most popular, as it includes a larger set of symptoms. We did not use it as one of the PROMs, feeling it was too long for routine clinical use in busy consultations at 21 questions, but we did choose to use it as our primary outcome measure.

Bryan and Margaret are members of the study group, and the MIND/Depression Alliance PPI advisory group in Southampton will meet every few months to review progress and give further advice. We plan that at least one PPI colleague will attend each study team meeting, so that we always have their support and input. Each will also be buddied with a named academic member of the study team, to act as first port of call for support (e.g. explanation of jargon), and liaison in between team meetings. They will be paid £18.75 per hour for their time, which will include attending study team meetings and commenting on relevant documents through email, plus travel and any other out-of-pocket expenses (in line with INVOLVE recommendations). Bryan has also agreed to help publicise the results of the trial through his Depression Alliance group locally, and nationally through Mind. We will give PPI colleagues regular feedback on our interactions with them, and ask for theirs.

PROJECT MANAGEMENT

Trial Steering Committee and Independent Data Monitoring Committee

A Trial Steering Committee (TSC), and Independent Data Monitoring Committee (IDMC) will be set up to oversee trial conduct. Adverse events will be reported to the IDMC who will advise the TSC about continuation and whether interim analyses are needed. The TSC will work with the IDMC and be kept informed by the CI, PI, or Trial Coordinator. If an extension was asked for then it would be the responsibility of the TSC to look in detail as to why this was needed and give an opinion which would inform the funder (NIHR) and the sponsor (University of Southampton).

Study group meetings

TK, CD, and GLew will lead weekly local study team meetings in Southampton, Liverpool and North London respectively, and an overall trial management group (TMG) will meet every month, through teleconferencing, to review progress and give advice on the conduct and management of the study.

TK, CD and GLew will recruit one and a half full-time researchers in each of the three centres and liaise directly with participating practices during recruitment and at intervals throughout the trial. The researchers will be responsible for carrying out the trial within the practices on a day-to-day basis, including consenting and randomising participants and assessing research outcomes, and will meet with TK, CD, or GLew weekly throughout the course of the study. The researchers will lead on data entry, supported by the administrative assistants. BS will lead on the quantitative analysis working with TK, MM, PL, CD, GLew, and CD. The researchers will complete the qualitative interviews of patients and GPs/NPs. GLey, CD, AG and CD will advise the researchers on the development of the semi-structured qualitative interviews and the iterative analysis. Heather Minett, Finance Officer for the Faculty of Medicine at the University of Southampton, will manage the budget and report on expenditure annually.

SAFETY OF PARTICIPANTS IN THE TRIAL

We have been advised by the Medicines and Healthcare Products Regulatory Authority (MHRA) that the study is not a clinical trial of an investigational medicinal product (CTIMP), and so a Clinical Trials Authorisation (CTA) is not required. Therefore the study team will not be bound by the MHRA CTIMP regulations on alerting the sponsor and ethics committee within specific timescales of any adverse events, or potential adverse effects of the intervention or trial procedures, which are reported by patients or practitioners participating in the trial.

However, the safety of patients in the trial remains our paramount consideration and the trial coordinator will ensure that any adverse events reported by patients or practitioners will be brought to the attention of the Programme Manager as quickly as possible and ideally, within 24hrs. It will then be for immediate discussion with the Chief Investigator (CI) or in the absence of the CI, one of the Principal Investigators (PIs). The CI or PI will decide whether or not to inform Sponsor or the Ethics Committee, PSC or IDMC. The report will include the event, when the information was reported, assessment of seriousness and likely relationship to participation in the trial.

All serious adverse events (SAEs) will be reported to the Chief Investigator and the Trial Coordinator within 24 hours of the local site becoming aware of the event. We will use the Southampton CTU's SAE Non-CTIMP Form, which asks for the nature of the event, date of onset, severity, corrective therapies given, outcome, causality (i.e. unrelated, unlikely, possible, probably, definitely) and expectedness. The Chief Investigator will assign the causality and expectedness of the event and the term should be in accordance with the latest version of MedDRA and grades given in accordance with the NCI CTCAE v4.03. Additional information will be provided as soon as possible if the event has not resolved at the time of reporting.

The Chief Investigator or Programme Manager will notify the REC of related and unexpected SAEs occurring during the study according to the following timelines; fatal and life-threatening within 7 days of notification and non-life threatening within 15 days.

WITHDRAWAL FROM THE STUDY

Patient and practitioner participants will be free to withdraw consent at any time without providing a reason. When withdrawn, patient participants will continue to receive standard clinical care from their practitioner. Follow up data will continue to be collected (unless the participant has specifically stated that they do not want this to happen).

CONFIDENTIALITY

The research team will preserve the confidentiality of participants taking part in the study. The investigators will ensure that participant's anonymity will be maintained and that their identities are protected from unauthorised parties. On trial documents and files participants will not be identified by their names, but by an identification code. The key to identification codes will be kept in a separate room to the trial data documents and files, in a locked cabinet within the University of Southampton Primary Care department at Aldermoor Health Centre.

INDEMNITY

The University of Southampton's public and professional indemnity insurance policy provides an indemnity to UoS employees for their potential liability for harm to participants during the conduct of the research. This does not in any way affect an NHS' Trust's or GP Practice's responsibility for any clinical negligence on the part of its staff.

DATA HANDLING

Participant data will be entered on laptop computers on site and retained at the University of Southampton in accordance with the General Data Protection Regulation (2018). The CI will be responsible for ensuring the accuracy, completeness, and timeliness of the data entered. Participant data will be pseudo-anonymised by assigning each participant a participant identifier code which will be used to identify the participant during the study and for any participant- specific clarification between the University and participating practices.

The Informed Consent Form will specify the participant data to be collected and how it will be managed or might be shared; including handling of all Patient Identifiable Data (PID) and sensitive PID adhering to relevant data protection law. Trained personnel with specific roles assigned will be granted access to the electronic patient data.

MONITORING

Data stored at the University of Southampton will be checked for missing or unusual values (range checks) and checked for consistency within participants over time. Any suspect data will be returned to the researcher or practice in the form of data queries.

The participants' medical records and other relevant data may also be reviewed by appropriate qualified personnel independent from the trial team, appointed to audit the study, including representatives of the Competent Authority. Details will remain confidential and participants' names will not be recorded outside the University.

RECORD RETENTION AND ARCHIVING

Study documents will be retained in a secure location during and after the trial has finished, in accordance with the sponsor University of Southampton's regulations. After study closure the CI will maintain all source documents and study related documents and retain them for a period of 10 years.

PUBLICATION AND DATA SHARING POLICY

A trial dissemination group will be established, whose purpose will be to oversee the planned outputs from the REDUCE programme, and agree on data sharing arrangements. This group will comprise the CI Tony Kendrick (TK) in Southampton and one co-applicant from each of the other two centres - Liverpool (Chris Dowrick), and London (Glyn Lewis).

REFERENCES

References

1. Jorm AF, Patten SB, Brugha TS, Mojtabai R. Has increased provision of treatment reduced the prevalence of common mental disorders? Review of the evidence from four countries. World Psychiatry 2017 doi:10.1002/wps.20388
2. NICE (National Institute for Health and Care Excellence). Depression. The treatment and management of depression in adults. CG90. Draft 2017 update published for consultation at: https://www.nice.org.uk/guidance/indevelopment/gid-cgwave0725
3. Kendrick T, King F, Albertella L, Smith PW. GP treatment decisions for patients with depression: an observational study. BJGP 2005; 55(513): 280–286.
4. Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. Lancet. 2009;374(9690):609-19. doi: 10.1016/S0140-6736(09)60879-5.
5. Dowrick C, Leydon GM, McBride A, et al. Patients' and doctors' views on depression severity questionnaires incentivised in UK quality and outcomes framework: qualitative study. BMJ 2009; 338(7697): 1–9.
6. Kendrick T, El-Gohary M, Stuart B, et al. Routine use of patient reported outcome measures (PROMs) for improving treatment of common mental health disorders in adults. Cochrane Database of Systematic Reviews 2016, Issue 7. Art. No.: CD011119. DOI: 10.1002/14651858.CD011119.pub2.
7. McManus S, Bebbington P, Jenkins R, Brugha T. (eds.) Mental health and wellbeing in England: Adult Psychiatric Morbidity Survey 2014. Leeds: NHS Digital, 2016.
8. McCrone P, Dhanasiri S, Patel A, Knapp M. Paying the price: the cost of mental health care in England to 2026. King's Fund, London, 2008.
9. Black N. Patient reported outcome measures could help transform healthcare. BMJ 2013;346:f167 doi: 10.1136/bmj.f167.
10. Kendrick T, Dowrick C, McBride A, et al. Management of depression in UK general practice in relation to scores on depression severity questionnaires: analysis of medical record data. BMJ 2009; 338:b750.
11. Moore M, Ali S, Stuart B, et al. Depression management in primary care: an observational study of management changes related to PHQ-9 score for depression monitoring. BJGP 2012; DOI:10.3399/bjgp12X649151.
12. Shaw EJ, Sutcliffe D, Lacey T, Stokes T. Assessing depression severity using the UK Quality and Outcomes Framework depression indicators: a systematic review. BJGP 2013; DOI:10.3399/ bjgp13X667169.
13. Kendrick, T., Stuart, B., Newell, C., et al. Did NICE guidelines and the Quality Outcomes Framework change GP antidepressant prescribing in England? Observational study with time trend analyses 2003-2013. J Affective Dis 2015 186, 171-177. DOI: 10.1016/j.jad.2015.06.052
14. Delgadillo J, Overend K, Lucock M, et al. Improving the efficiency of psychological treatment using outcome feedback technology. Behav Res Ther 2017 99:89-97. doi: 10.1016/j.brat.2017.09.011.
15. Health & Social Care Information Centre. Psychological Therapies, Annual Report on the use of IAPT services: England– 2013/14. http://www.hscic.gov.uk/catalogue/PUB14899
16. Knaup C, Koesters M, Schoefer D, Becker T, Puschner B. Effect of feedback of treatment outcome in specialist mental healthcare: meta-analysis. Br J Psychiatry (2009) 195, 15–22. doi: 10.1192/bjp.bp.108.053967.
17. Shimokawa, K., Lambert, M. J., Smart, D. W. Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. Journal of Consulting and Clinical Psychology, 2010;78:298 –311. http://dx.doi.org/10.1037/a0019247
18. Kendrick, T., Stuart, B., Leydon, G. M., et al. Patient-reported outcome measures for monitoring primary care patients with depression: PROMDEP feasibility randomised trial. BMJ Open 2017, 7(3), 1-11. [e015266]. DOI: 10.1136/bmjopen-2016-015266.
19. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression measure. J Gen Intern Med 2001;16:606-613.

20. Roth, A. J., Kornblith, A. B., Batel-Copel, L., et al. Rapid screening for psychologic distress in men with prostate carcinoma. Cancer 1998;82:1904–1908. doi:10.1002/(SICI)1097-0142(19980515)82:10<1904:AID-CNCR13>3.0.CO;2-X.
21. Ashworth M, Kordowicz M, Schofield S. 'PSYCHLOPS' (Psychological Outcome Profiles): an outcome measure. Integrating Science and Practice 2012;2:36-39.
22. Yeung AS, Jing Y, Brenneman SK, et al. Clinical Outcomes in Measurement based Treatment (COMET): a trial of depression monitoring and feedback to primary care physicians. Depress Anxiety 2012; 29(10): 865–873.
23. Chang TE, Jing Y, Yeung AS, et al. Effect of communicating depression severity on physician prescribing patterns: findings from the Clinical Outcomes in Measurement-based Treatment (COMET) trial. Gen Hosp Psychiatry 2012; 34(2): 105–112.
24. Pettersson A, Björkelund C, Petersson EL. To score or not to score: a qualitative study on GPs' views on the use of instruments for depression. Fam Pract 2014;31:215–21. doi.org/10.1093/fampra/cmt082
25. Wikberg C, Westman J, Petersson EL, et al. Use of a self-rating scale to monitor depression severity in recurrent GP consultations in primary care – does it really make a difference? A randomised controlled study. BMC Fam Practice (2017) 18:6 DOI 10.1186/s12875-016-0578-9
26. May CR. Normalization Process Theory. http://www.normalizationprocess.org/
27. Spitzer RL, Williams JBW, Kroenke K. Instructions for Patient Health Questionnaire (PHQ) measure. https://phqscreeners.pfizer.edrupalgardens.com/sites/g/files/g10016261/f/201412/instructions.pdf (accessed 17th January 2018).
28. Löwe B, Blankenberg S, Wegscheider K et al. Depression screening with patient-targeted feedback in cardiology: DEPSCREEN-INFO randomised clinical trial. Br J Psychiatry 2017 doi: 10.1192/ bjp.bp.116.184168
29. Reeve BB, Wyrwich KW, Wu AW, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. Qual Life Res 2013;22:1889–905. http://dx.doi.org/ 10.1007/s11136-012-0344-y
30. Gilbody S, Richards D, Barkham M. Diagnosing depression in primary care using self-completed instruments: UK validation of PHQ–9 and CORE–OM. BJGP 2007;57:650-52.
31. Löwe B, Kroenke K, Herzog W Grafe K. Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). J Affective Dis 81 (2004) 61–66.
32. Fitzpatrick R, Gibbons E, Mackintosh A. *An overview of patient-reported outcome measures for people with anxiety and depression.* Oxford: Patient-reported Outcome Measurement Group, Department of Public Health, University of Oxford, 2009.
33. Cameron IM, Crawford JR, Lawton K, Reid IC. Psychometric comparison of PHQ-9 and HADS for measuring depression severity in primary care. Br J Gen Pract. 2008; 58:32-6. doi: 10.3399/bjgp08X263794.
34. Robinson J, Khan N, Fusco L, et al. Why are there discrepancies between depressed patients' Global Rating of Change and scores on the Patient Health Questionnaire depression module? A qualitative study of primary care in England. BMJ Open. 2017 4;7(4):e014519.
35. Michie S, Van Stralen M, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implementation Science 2011;6:42 doi:10.1186/1748-5908-6-42.
36. Beck AT, Steer RA, Brown GK. Manual for the Beck Depression Inventory-II. Psychological Corporation, San Antonio, Texas, 1996.
37. Spitzer RL, Kroenke K, Williams JB, et al; A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006;166:1092-7.
38. Mundt JC, Marks IM, Shear MK, et al. The work and social adjustment scale: a simple measure of impairment in functioning. Br J Psychiatry 2002;180:461–4.
39. EuroQol Group (1990). EuroQol-a new facility for the measurement of health-related quality of life. Health Policy 16(3):199-208.
40. Beecham J, Knapp, M. Costing psychiatric interventions. In: Measuring Mental Health Needs (eds. Thornicroft G, Brewin C, Wing J), pp. 163–183. Gaskell, London, 1992.

41. Meakin R, Weinman J. The 'Medical Interview Satisfaction Scale' (MISS-21) adapted for British general practice. Fam Pract 2002 Jun;19(3):257-63.
42. Vanheule S, Desmet M, Groenvynck H, et al. The Factor Structure of the Beck Depression Inventory–II. An Evaluation. Assessment 2008;15:177-187. doi/10.1177/ 1073191107311261.
43. Dolan P. Modelling variations for EuroQol health states. Medical Care 1997;35:1095-1108.
44. Button KS, Kounali D, Thomas L, et al. Minimal clinically important difference on the Beck Depression Inventory – II according to the patient's perspective. Psychol Med 2015;45:3269–3279 doi:10.1017/S0033291715001270.
45. Murray E, Treweek S, Pope C, MacFarlane A, Ballini L, et al. Normalisation process theory: a framework for developing, evaluating and implementing complex interventions. BMC Med 2010;8:63.
46. Glaser, B. The constant comparative method of qualitative analysis. Social Problems 1965;12: 436–445.
47. Curtis, L. & Burns, A. *Unit Costs of Health and Social Care 2017.* Personal Social Services Research Unit, University of Kent, Canterbury. https://doi.org/10.22024/UniKent/01.02/65559