Study Protocol

The Feasibility and Implementation of a Psychosis Risk Prediction Algorithm (P Risk) for use in Primary Care

Protocol for the qualitative work package

Protocol version and date: version 0.7 20/04/23

IRAS number: 315320

Chief Investigator and co-applicants: Sarah Sullivan (University of Bristol), David Kessler (University of Bristol), Katrina Turner (University of Bristol), Gareth Ambler (University College London), Irwin Nazareth (University College London), Harminder Kaur (Patient and Public Contributor).

LIST OF CONTENTS

KEY STUDY CONTACTS

Chief Investigator	Dr Sarah Sullivan, sarah.sullivan@bristol.ac.uk
Sponsor	University of Bristol, research-governance@bristol.ac.uk
Funder(s)	1] NIHR RfPB, Richard Foulsham, Richard.foulsham@nihr.ac.uk 2] Biomedical Research Centre, jack.teasdale@bristol.ac.uk

STUDY SUMMARY

Study Title	The feasibility and implementation of a psychosis risk prediction algorithm (**P Risk**) for use in primary care.
Short Title	Psychosis risk prediction in primary care
Study Participants	Service users, carers, GPs and EIT practitioners
Planned Size of Sample	10-12 GPs, 6-8 EIT practitioners, 16-18 service users, 10-12 carers
Planned Study Duration	14 months
End of Study Definition	When all data are collected and analyses are complete
Research Question	Is it feasible and acceptable to practitioners, patients, and carers to implement a psychosis risk prediction algorithm in primary care?

ROLE OF STUDY SPONSOR AND FUNDER

The study Sponsor takes ultimate responsibility for the oversight of the research project. For more detail on the role of the Sponsor, please see: http://www.bris.ac.uk/red/research-governance/ethics/sponsorship/

The funder (NIHR) has had influence on the study design resulting from the comments of peer reviewers but will have no influence on the conduct, data analysis, interpretation, manuscript writing or dissemination of the results. The second funder (NIHR BRC) has had/will have no influence on the study design, conduct, data analysis, interpretation, manuscript writing or dissemination of the results.

ROLES AND RESPONSIBILITIES OF THE STUDY MANAGEMENT GROUP

The study management group (SMG) will consist of the entire study team (all co-applicants, the PPI members, and the study researcher). The SMG will meet at least monthly throughout the study duration. It will be responsible for ensuring that the study meets its aims to time and budget. The SMG is independent from the Sponsor. The Patient and Public Involvement Group will consist of the Chief Investigator and the two PPI members of the study team. This group will also meet at least monthly. They will help to write the topic guide for the focus groups and assist with any problems with recruitment for these groups.

STUDY FLOWCHART

	Study Timeline		Aug	Sept	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept
		Year	2022					2023								
Task	**Work Package One**															
	Select PCN network															
	Implement P risk algorithm in EMIS software															
	Run P risk on historical records															
	Problem solving for software bugs															
	Work Package Two															
	Link primary and secondary care data															
	Collect data on number of referrals															
	Collect data on number of EIT assessments															
	Collect data on diagnoses															
	Collect data on diagnoses but not referred by GP															
	Work Package Three															
	Assemble cohort sample															
	Apply P risk to cohort															
	Statistical analysis for cohort															
	Assemble nested case control sub-samples															
	Apply P risk to case control sub-samples															
	Statistical analysis for case control sub-samples															
	Work Package Four															
	Purposefully sample GP practices															
	Identify EITs															
	Purposefully sample patients															
	Design topic guides															
	Conduct GP interviews															
	Conduct EIT interviews															
	Conduct patients focus groups															
	Analyse qualitative data															
	Dissemination															
	Publish paper on quantitative findings															
	Publish paper on qualitative findings															
	Study wrap up															

STUDY PROTOCOL: The Feasibility and Implementation of a Psychosis Risk Prediction Algorithm (P Risk) for use in Primary Care.

1. Background and Rationale

Psychosis is a term for a group of serious, long-term mental illnesses characterised by loss of contact with reality, e.g. schizophrenia. The World Health Organisation ranks psychosis as the third most disabling condition worldwide[1] and in 2013 there were 23.6 million prevalent cases worldwide[2]. Psychosis has a lifetime prevalence of 3%[3] and peak incidence is in young adulthood[4].

Psychosis is associated with a significant global disease burden[5]. It is the most expensive brain related disorder and is more expensive than cardiovascular disease[5], which has a much higher prevalence. An analysis in England reported annual direct costs of £2 billion and indirect costs of £4.7 billion[6].

Most people with psychosis in the UK enter specialist mental health care (i.e. Early Intervention for Psychosis Teams – EITs) via a referral from their GP[7] and a shorter duration of untreated psychosis (DUP) is associated with more general practice (GP) visits prior to diagnosis date[8]. GPs therefore are a vital part of the psychosis care pathway[8]. It is consequently important that GPs recognise the warning signs of psychosis to expedite referral to EITs, but this is difficult to achieve. The accuracy of psychosis diagnoses recorded in primary care electronic health records (EHRs) is high[9]; however there can be a delay because the early symptoms that are precursors to psychosis are common[10] and not obviously identified with psychosis e.g. sleep disruption. Delays in GP referral occur because: 1] nonpsychotic warning symptoms of psychosis are non-specific. 2] most GPs do not develop diagnostic skills for the warning signs of psychosis because, individually, they see only a few cases. 3] patients do not always have continuity of care and may see different GPs, meaning that subtle changes in their mental state, e.g. increasing anxiety, are missed. There is potential to help GPs identify the early warning signs of psychosis if the data already in EHRs could be used to alert them to the possibility of early psychosis. Data in EHRs could help build a prediction algorithm that is more sensitive to patterns of symptoms and the weightings between them, than would be possible in a traditional primary care clinical situation. In this way, data-driven algorithms have the potential to support diagnostic acumen.

A previous nested case (n=11,690) control (n=81,793) study[11] in primary care EHRs found that 12 predefined nonpsychotic warning symptoms and signs (depression, obsessive compulsive disorder, attention deficit hyperactivity disorder, mania, blunted affect, problems with cannabis, problems with cigarette smoking, sleeping problems, suicidal behaviour, bizarre behaviour, social isolation, role functioning problems), sociodemographic factors and consultation frequency were associated with

later psychosis. The study findings suggested that these factors could be candidate predictors for a risk prediction model for new onset of psychosis.

Using these findings we have developed and internally validated a prediction model (**P Risk**), using a sample of 300,000 people from 216 GP practices with linked primary care and secondary care EHRs and at least 5 years of follow up data, who had consulted their GP for any nonpsychotic mental health problem (paper submitted for publication May 2022). From this sample 830 diagnoses of psychosis were detected. We used robust multivariable estimation with shrinkage (to avoid overfitting) for variable selection during model development and internal validation used bootstrapping methods. **P Risk** discrimination (i.e. the ability to discriminate between those who later develop psychosis and those do not) performance is good[12] (Harrell's C statistic of 0.77). Once the final model had been identified, we calculated the predicted risk over 6 years for patients whose follow up time exceeded 6 years, or who had experienced an event within 6 years. We calculated sensitivity, specificity, and likelihood ratios for thresholds of risk of 0.5%, 1%, 1.5% and 2%. The sensitivity (68.9%) and specificity (70.9%) were similar and high for a 1% risk threshold, the likelihood ratio was 2.37.

External validation of **P Risk** is ongoing in a separate EHR database and is expected to show similar performance because the methods used in model development reduced overfitting. This work will be complete by the end of May 2022. The external validation work will produce a psychosis risk probability threshold for low, medium, and high risk of psychosis.

P Risk overcomes the problems that GPs have in identifying patients with the early warning signs of psychosis described above by using existing primary care data to predict future risk. It uses EHR data on GP consultations for 14 proven nonpsychotic predictors for psychosis. It can therefore support individual GP diagnostic skills and is not dependent on care continuity. **P Risk** could be automated for use on GP IT systems in a similar way to other tools already in use, such as Q Risk for detecting cardiac risk and FRAX for osteoporosis risk. This would be the first example of a psychosis primary care prediction tool and there is evidence that automated risk predictors can improve patient outcomes[13].

Although we have evidence that **P Risk** is statistically accurate, a feasibility and implementation study is required to investigate its operationalisation and acceptability in real-world clinical situations and to inform the design and delivery of a future evaluation to establish effectiveness and cost-effectiveness. Secondary care models have been developed to predict transition to psychosis in

individuals considered to be at clinical high risk of psychosis (CHR-P) or who are already receiving mental health treatment. However, the findings of a scoping review, indicated that there are currently no psychosis prediction models that have been developed and validated for use in primary care. Also, a recent meta-analysis of studies of transition to psychosis in CHR-P secondary-care patients found 7 which described prediction models. These results are encouraging, but the studies described models with predictors that are not routinely collected in primary care (i.e. results from complex cognitive tests and brain scans).

2. Theoretical Framework

The study will be pragmatic in nature but informed by grounded theory[14] and the person-based approached to intervention development[15]. Both approaches will ensure that as data collection and analysis progress, the insights gained inform sampling of participants and ongoing data collection.

3. Research Question and Aims and Objectives

Research question: Is it feasible and acceptable to practitioners, patients and carers to implement a psychosis risk prediction algorithm in primary care.

Aims and Objectives

Overall Aim: To determine the operationalisation and acceptability of using **P Risk** in real-world clinical situations.

Objectives:

Work package One: Demonstrate whether **P Risk** can be implemented in primary care data systems.

Work package Three: Investigate the accuracy of **P Risk** using real-world primary care data systems

Work package Four: Investigate the acceptability of **P Risk** to practitioners, patients, and carers.

We are only seeking NHS REC and HRA approval for Work package four. In addition, we are seeking University of Bristol approval for work package Three. Work package One and Three do not require NHS REC and HRA approval.

4. Study Design and Methods of Data Collection and Data Analysis

Work package 1:

Study design

Conversion of the P Risk statistical algorithm into a code that is compatible with EMIS medical records software. Detecting any bugs in the code and correcting these.

Population

One Primary Care Network (PCN) in Bristol, North Somerset and South Gloucestershire CCG (BNSSGCCG) region.

Sample Size

4 GP practices with approximately 64,000 patients registered. This is likely to lead to an estimated 400 referrals to EITs for a psychosis referral and 147 diagnoses of a psychotic disorder.

Inclusion/Exclusion Criteria

Inclusion: GP practices which use EMIS clinical records software. All GP practices in the BNSSG region use EMIS software. Exclusion: None.

Datasets

Anonymised electronic medical records of all patients in 4 PCNs who have at least 5 years of previous consultation history.

Procedure

EMIS Connecting Healthcare plc, a study collaborator, have experience of implementing risk prediction tools e.g. Q Risk. Using this experience and the algorithmic information provided by the study team, EMIS will remotely install **P Risk** on the PCN IT systems of the study practices, utilising software engineering functions.

The **P Risk** algorithm will be run on the historical EHRs of every patient over the previous 5 years to investigate any 'bugs' in the functioning of **P Risk** in EMIS software. Data will be collected on the software 'bugs' detected, including any solutions found during software engineering problem solving exercises. These solutions will be useful when extending the use of **P Risk** to other primary care software such as System One.

Work package 3:

Population: All BNSSGCCG GP practices.
Inclusion criteria: GP practices which use EMIS clinical records software.
Exclusion criteria: Any patient with an existing coded diagnosis of psychosis either in primary or secondary care EHRs or any recorded prescription for anti-psychotic medication at a dosage

appropriate for psychosis. This is to ensure that incident, rather than prevalent, outcomes are detected by P Risk.

Sample size: Approximately 1 million patients whose GP electronic health records are linked to secondary care mental health electronic health records. We estimate that this sample will yield approximately 5000 incident EIT referrals and 2300 incident diagnoses of psychosis.

Designs: 1] Retrospective cohort and 2] nested case control studies.

Data Linkage: In the BNSSG CCG region OneCare will link primary and secondary care data and BNSSG CCG will supply a funded researcher to conduct the linkage work before the study commences.

Cohort Study Procedure

We will follow up all patients in the sample from the first recorded GP consultation for any reason recorded in primary care EHRs until one of the following occurs: death, the patient leaving the GP practice or the outcome (defined below).

Outcome: Coded incident diagnosis of First Episode Psychosis (FEP) or an At-Risk Mental State (ARMS) recorded in primary and/or secondary care EHRs.

Predictor: P Risk score. We will apply P Risk to the historical EHRs of the entire sample to establish the risk probability for each patient.

Data Analysis: We will calculate the absolute risk, calibration by calculating the C statistic, an Area under the Curve analysis, a calibration slope and calibration in the large. We will calculate the positive predictive value of P Risk (A/A+B) (see Figure 1).

Nested Case Control Study Procedure

We will investigate differences in P Risk accuracy in two subsamples of patients where there is evidence that they are at increased risk. These sub-groups are 1] patients of Afro-Caribbean ethnicity and 2] older women (50-65 years of age).

Sample: 1] Cases: Recorded Afro-Caribbean ethnicity AND a coded diagnosis of First Episode Psychosis (FEP) or an At-Risk Mental State (ARMS) recorded in primary and/or secondary care EHRs. Controls: Randomly selected at a ratio of at least 1:1 from the cohort without the outcome and of a non Afro-Caribbean ethnicity. 2] Cases: Female, aged between 50 and 65 years of age. Controls: Randomly selected at a ratio of at least 1:1 from the cohort without the outcome and women and men under 50-65 years of age. If more statistical power is required for either subgroup analysis the ratio of cases to controls can be increased up to 1:7.

Outcome: Coded diagnosis of First Episode Psychosis (FEP) or an At-Risk Mental State (ARMS) recorded in primary and/or secondary care EHRs.

Predictor: P Risk score. We will apply P Risk to the historical EHRs of the entire sample to establish the risk probability for each patient and whether the risk threshold for high, medium or low risk (to be

determined during external validation) would have been triggered. Any inappropriate risk threshold triggers will be investigated.

Data Analysis: Investigate the accuracy of P Risk in real-world data by calculating the sensitivity (true positives) (A/A+C) specificity (true negatives) (D/B+D) of P Risk (number of diagnoses of psychosis correctly picked up by P Risk) against the gold standard of a coded psychosis diagnosis (Figure One). We will also investigate:

1] whether the P Risk threshold for high, medium or low is optimal

2] calibration in each sub-group

Figure One. Accuracy of P Risk

P Risk threshold exceeded	Coded diagnosis of psychosis	
	Y	N
Y	A	B
N	C	D

Work package 4: Qualitative acceptability study

Study design

In-depth interviews will be held with potential users of **P Risk** (GPs) and focus groups will be held with those on whom it would be used and those affected by it (patients and their carers), to explore their views on the acceptability and potential value and implications of using **P Risk** in general practice. We will also offer to arrange individual interviews for patients and carers who are interested in taking part in the study but who are unable to make any of the suggested focus group dates and times. As **P Risk** may alter referrals from general practice to Early Intervention Teams (EITs), interviews will also be held with EIT staff to assess their views of **P Risk**, and their thoughts about GPs making referral decisions informed by **P Risk** and whether they would accept referrals on this basis.

Study Settings

CRNs covering Bristol and London will send information about the study to practice managers in GP practices that vary in size and the socio-demographic characteristics of their patient populations. The CRNs will ask practice managers if they would be willing to help recruit GPs for interview. The CRNs will inform the research team of which practices are willing to support the study and how.

GPs and patients will be informed about the study via GP practices. Up to 12 GP practices will be purposefully sampled and recruited in and around Bristol and London. They will vary in size and the socio-demographic characteristics of their patient populations and will be informed about the study via their local Clinical Research Networks. Practices will be asked if they would be willing to support

the study by providing their GPs with information about the study and/or emailing or posting information to patients identified through screening medical records as having consulted a GP about nonpsychotic mental health problems over the previous 6 months.

In terms of recruiting EIT staff for interview, the research team will identify EITs (in Bristol and in London), which vary in terms of the socio-demographic characteristics of the populations that they serve. The research team will invite staff by email (see initial email EIT version 1.3 20.04.2023) who conduct psychosis assessments within EITs to take part in a video call interview.

Participants

We will interview by video call:

- Between 10-12 GPs working in different study practices, who vary in age, gender, and whether they view themselves as having a particular interest in mental health will be sampled and invited for interview by email (see initial email GP vs 1.3 20.04.2023) and GP information sheet vs1.6 20.04.2023). This process will be organised via the GP Practice Manager.
- Between 6-8 EIT staff who conduct psychosis assessments will be purposefully sampled across the EITs supporting the study, to ensure variation in job role, gender, ethnicity and years since qualification. They will be invited to participate by email (see initial email EIT vs 1.3 20.04.2023 and EIT information sheet vs 1.6 20.04.2023).

We will also conduct online focus groups/individual interviews with:

- Between 16-32 patients in up to 4 focus groups with between 4 and 8 participants in each. We will aim to include patients who reflect the characteristics of individuals on whom the tool would most likely be used in clinical practice i.e. patients who have consulted their GP about nonpsychotic mental health problems over the previous 6 months. The inclusion criteria of poor mental health will match the mental health predictors in the P Risk model, which are depression/anxiety, suicidal behaviour (including self-harm), attention deficit hyperactivity disorder, obsessive compulsive disorder, sleep disturbance, and substance abuse. Eligible patients will be sent a text with a link to the P Risk study inviting them to take part in the interview (https://prisk.blogs.bristol.ac.uk/home/patients/), or will be emailed/posted a study invite letter/Expression of Interest (EoI) form and information sheet from the practice (see initial patient invite letter/EoI vs 1.3 20.04.2023 and patient information sheet vs 1.6 20.04.2023). We will also offer to arrange individual interviews with patients who are interested in taking part but who are unable to make any of the suggested focus group dates and times.

- Between 8 and 16 carers in between 2-4 focus groups with between 4 and 8 participants in each. Carer participants will be identified by their friend or relative, who is already a participant in the P Risk study (see carer initial invite letter/EoI vs 1.3 20.04.2023), and carer information sheet vs 1.6 20.04.2023).
- Exclusion Criteria: Inability to provide informed consent.

Recruitment of GPs

We will recruit GPs through GP practices. The research team will send practice managers supportive of the study the invite email and information sheet and ask them to email these to their GPs (see initial email GP vs 1.3 20.04.2023) and GP information sheet vs 1.6 20.04.2023). The email sent to GPs will state that, if the individual is willing to be interviewed, they should advise the practice manager or researcher with information on their gender, age, ethnicity and whether they have a particular interest in mental health. This information will be used to purposefully sample GPs to ensure maximum variation among interviewees according to these variables, in addition to sampling across all study practices.

Once responses have been received and GPs sampled, the researcher will contact the GPs or practice managers to arrange an interview time. The interviews will be held on a virtual meeting platform, such as Teams or Zoom. Any GPs who indicate that they are willing to be interviewed, but who are not sampled for interview, will be emailed, and thanked for their support. At the beginning of the interview GP consent will be taken verbally by the researcher.

Recruitment of patients

GP practices who indicate they are willing to screen their medical notes to identify patients who have consulted their GP about nonpsychotic mental health problems over the previous 6 months, will email eligible individuals an invite letter/EoI form which asks patients for information on their age, sex and ethnicity and an information sheet (see patient invite email/letter/EoI form vs 1.3 20.04.2023 and patient information sheet vs 1.6 20.04.2023). Where the practice does not have an individual's email address, these documents will be posted. In addition, GP practices will also have the option of sending patients a text with a link to the P Risk study inviting them to take part in the focus group/interview (https://prisk.blogs.bristol.ac.uk/home/patients/). Individuals willing to participate will be asked to complete the EoI form and return this to the research team via email or a SAE that will be enclosed with the study invite. As some patients may struggle to fill in the EoI form electronically, they will also be given the option of providing the researcher with information on their age, sex and ethnicity in an

email text, and this will count as patients having expressed interest in taking part in the study. This information will be used to purposefully sample individuals. Reminder email invites (see patient reminder email vs 1.3 20.04.2023) (along with an information sheet) will be sent by practices two weeks after the initial invite, if there is no response.

Once responses have been received and patients sampled, the researcher will contact the individual to arrange a focus group time and to answer any questions they have. The focus groups will be held by video call. Any individual who indicates they are willing to participate, but who are not sampled, will be either telephoned or emailed to thank them for their support. We will also offer to arrange individual interviews for individuals who are interested in taking part but who are unable to make any of the suggested focus group dates and times.

At the end of the focus group/interview, the researcher will explain to patients that the study team is also interested in conducting focus groups/interviews with their carers, and will ask patients to fill in a form to indicate if their carer (a family member or friend) would also be interested in participating in a separate focus group/individual interview, and provide us with their contact details. Patients who struggle to fill in the form electronically, will be given the option of providing this information in a text email directly to the researcher.

Recruitment of carers

If a form is returned by a patient indicating that they have a carer who is interested in participating, the researcher will email or post (if no email address is provided) the carer an initial invite/EoI (see carer invite email/letter vs 1.3 20.04.2023) and carer information sheet (see carer information sheet vs 1.6 20.04.2023). Researchers will also have the option of texting the carers a link to the study, if patients indicate that this is their carers' preferred method of contact. Those who express an interest in taking part will be contacted by the researcher to arrange focus group dates and times. If no reply is received to the initial invite after two weeks, a reminder email/text will be sent (see carer reminder email vs 1.3 20.04.2023) along with a carer information sheet. We will also offer to arrange individual interviews for carers who are interested in taking part but who are unable to make any of the suggested focus group dates and times.

Recruitment of EIT practitioners

Clinical leads for the EIT services supporting the study will email a letter of invitation and an information sheet about the study to their practitioners (see EIT initial invite vs 1.3 20.04.2023 and

EIT information sheet vs 1.6 20.04.2023). The research team will also offer to join a team meeting remotely to talk about the study and answer any questions. EIT practitioners who are willing to take part will be asked to email the researcher or to let their clinical lead know they are interested (in this situation the clinical lead would inform the research team of their interest). The researcher will then contact the practitioner to answer any further questions and to arrange an interview time. The interview will be held on a virtual meeting platform and verbal consent taken at the time of interview.

Consent

Everyone invited to take part in the study will receive an information sheet that details the aims of the study, what the interviews/focus group will entail, how data will be stored and analysed. This information sheet will also state that the individual does not need to take part and if they do, they have the right to withdraw from the study at any point without giving any reasons. For all interviewees, at the time of interview/focus group, the individual will have time to ask any questions they may have. Verbal consent will be obtained from practitioners, patients and carers before the interview starts. This will consist of the interviewer reading aloud the consent form, and asking for participants' permission to intial and date the consent form. The process of consenting practitioners, patients and carers to the study will be audio-recorded separately from the qualitative interview, and will be stored separately from the research data. Only members of the research team will have access to the consent forms. At the end of the interview, the researcher will send the participant a copy of the consent form.

Data Collection: the interviews and focus groups

Topic guides will be used during data collection to ensure consistency across the interviews and focus groups. Four guides will be developed (GP, EIT practitioner, patient, and carer). They will be based on the aims of the interviews/focus groups, relevant literature, and discussion amongst team members. They will be developed in parallel to ensure issues relevant to each interviewee group are explored, enabling comparison of views across the four data sets during data analysis.

The interviews and focus groups will be conducted by a Research Associate experienced in conducting and analysing qualitative interviews and focus groups. With participant consent, the interviews/focus groups will be audio recorded using an encrypted solid state voice recorder/the virtual platform and transcribed verbatim by a University of Bristol approved professional transcribing service. The study team signed a confidentiality agreement with the transcription service that we will be using as part of the study.

The purpose of the GP interviews will be to explore their views of the value and use of the prediction tool, how they think it should be embedded in clinical practice and results communicated to them and their patients, and how this would affect the process of referring to secondary care for a psychosis assessment. We will also ask their opinions of what might hinder and facilitate the use of **P Risk** in practice, what the referral thresholds should be, how results should be presented and communicated to GPs, and between services, and the extent to which **P Risk** results would be used to inform their decisions to refer/accept patients into EIT services. We will also ask GPs for their views on conducting a trial to investigate the effectiveness of **P Risk** and their general willingness to engage in mental health research like this. GPs will be interviewed on a virtual platform such as Zoom or Teams to encourage participation, and because video-call interviews can gather the same information as those conducted in person[16, 17]. The interviews will last up to 30 minutes. GPs practices will be reimbursed for each of their GP interviewed.

The interviews held with EIT practitioners will focus on their views of the prediction tool, how GPs should communicate results to patients and practitioners the individual would be referred to, what referral thresholds should be, what might hinder and facilitate the use of P Risk in practice, and how they think patients identified as being at risk should be managed and supported. We will also ask practitioners for their views on conducting a trial to investigate the effectiveness of **P Risk** and their general willingness to engage in mental health research like this. These interviews will be held on a virtual platform such as Zoom or Teams and will last up to 30 minutes. EITs will be reimbursed for each of their clinicians interviewed.

The focus groups with patients and carers will focus on their views of using a data driven algorithm to provide a risk prediction and how they would like this prediction communicated to patients and carers. We will also ask participants for their views on conducting a trial to investigate the effectiveness of **P Risk** and their general willingness to engage in mental health research like this. The focus groups will last up to one hour. Participants will receive incentives to compensate them for time taken during focus groups/individual interviews.

Data Analysis

When conducting the interviews and focus groups, data collection and analysis will proceed in parallel, so that analytical insights from the initial interviews/focus groups can shape later data collection and to enable the team to establish when data saturation has been reached. The interviews and focus

groups will be audio-recorded, transcribed verbatim and anonymised. All the data will be analysed thematically so that comparisons can be made within and across the data sets.

Two researchers will independently read and code a sub-sample of interviews and focus groups from the four data sets. The codes and themes identified will be discussed and four preliminary coding frames developed. Where possible, similar codes will be used within these coding frames to assist comparisons across the different participant groups. After coding frames have been agreed by all members of the research team, transcripts will be uploaded to the software package NVivo to aid data management and analysis. Each transcript will then be coded electronically. Once this has been done, using an approach based on Framework analysis[18], data coded under a specific code will be electronically retrieved using NVivo, and then summarised in a table formatted so that the rows represented each participant and columns relevant codes.

Initially, data collected from GPs, EIT practitioners, patients and carers will be analysed separately to ensure detailed understanding of each groups' accounts is gained before comparisons are made between them. Findings from the four groups of interviewees will then be triangulated to provide a fuller understanding of the areas explored with each group of participants, and to increase the confidence with which conclusions can be drawn.

During the focus groups, with the consent of participants, discussions will be video and audio recorded for the purpose of ensuring accurate notetaking. Following the focus groups, these notes will be read and re-read by the research team and a table created, listing key points and how each will be addressed.

Data Management and Storage

Completed consent forms and audio files of the consent process will be kept separately from the research data (i.e. research interview transcripts and research audio sound files) on a project folder created by the qualitative researcher and saved short to medium-term on secure servers of the Bristol Medical School. The qualitative interview sound files will be downloaded to the study folder immediately after interview/focus group and labelled with a unique identification (ID) number. In the terms of the GDPR this makes it pseudonymised data. The research team will complete a DPIA checklist and a full DPIA if required. Data will then be deleted from the solid-state voice recorder, and then deleted from the project folder once transcripts have been checked for accuracy. Transcripts will be labelled according to the participant's/group's unique ID number and anonymised. A list that

matches participants' names and contact details, with their ID number will be kept on a password protected document in a different study folder.

When consenting to take part in the study, interviewees/focus group attenders will be asked to consent to their anonymised transcript being stored for use in future research studies. Each week, the study folder will be replicated by the PI on the Bristol University's Research Data Storage Facility (www.acrc.bris.ac.uk) for long-term storage. Both the Medical School's server and the Research Data Storage Facility (RDSF) are regularly backed up. The data controller will be the University of Bristol. Data will be stored in the RDSF but made available to other researchers via Bristol University's Research Data Repository (https://data.bris.ac.uk/data/). Each deposit is accompanied by appropriate metadata and is assigned a unique Digital Object Identifier (DOI) via the DataCite scheme. All DataCite DOIs are searchable, internationally. Each deposit has a 'readme.txt' that details the content of the data set and provides the DOI of any journal publications from the project. These journal publications will also be useful to other users, as they will detail the methods and procedures used to collect and analyse the data.

The lawfulness, fairness and transparency principle of GDPR requires data subjects to be provided with certain information about the collection and processing of their personal data. This information is typically provided in the form of 'fair processing notices' or 'privacy notices'. In a research context, this information may often be included within a patient information sheet or similar document which is provided to support informed consent, though it can be provided separately. The GDPR require the information to be provided to data subjects in a concise, transparent, intelligible and easily accessible form and to be written in clear and plain language (in particular where directed to a child).

Ethical and Regulatory Considerations

Assessment and management of risk

Interviews will be held by a researcher skilled in qualitative research. It is unlikely any sensitive or upsetting material will be disclosed. However, if any interviewee makes a comment that suggests they are a risk to themselves or others, the researcher will discuss this with them and explain that the individual's GP will be notified of this.

Amendments

Protocol amendments will only be implemented when approved by the NHS Research Ethics Committee and HRA, as appropriate. The study's sponsor will review and approve any amendments prior to submission.

Peer Review

This study is funded by the NIHR and therefore has been externally reviewed by other academics and members of a funding committee. Approval is being sought from NHS REC and HRA.

Patient and Public Involvement

Application Development

The research question originated from a series of psychosis service user (SU) forum events between 2017-2019. At these events SUs described problems they had experienced getting their GP to recognise their warning symptoms of psychosis. In January 2021 we held two virtual PPI events with 13 attendees from across the UK, with lived experience of psychosis or of caring for someone with psychosis. We asked for their views on the proposed study, and their input has directly informed development of this proposal e.g. empathetic ways that the psychosis risk prediction can be communicated by clinicians and the importance of depression and self-harm as predictors. From this group we recruited a Lived Experience Advisory Panel (LEAP) which has both SU and carer representation. LEAP is facilitated by JR and HK and has so far met three times to discuss proposal development. During these meetings we have provided bespoke training for the PPI co-applicants and the LEAP group and plan further training. Two people (HK and JR) with lived experience of psychosis are also co-applicants on this proposal. JR will act as a mentor for the LEAP and will be supported by the PI. The LEAP are remunerated at INVOLVE rates. A LEAP made up of African Caribbean members has also been convened, who will also be consulted throughout the study.

Research Involvement

The LEAP will meet every two months throughout the study and be consulted by the study team to ensure that the study remains relevant to SU/carer concerns. The African Caribbean LEAP member (see above) who have been recruited, will be invited to join the study LEAP. The LEAP will: 1] comment on and edit SU/carer facing materials. 2] help write the topic guides and interpret qualitative data. 3] help facilitate dissemination of study findings to key stakeholder groups. This involvement will be managed by the PPI co-applicants.

Data Protection and Patient Confidentiality

Each participant/group will be assigned a unique Identification (ID) number which will help researchers to identify participants throughout the study without disclosing their identity. ID numbers will replace personal identifiable data in all research related documents. Any personally identifiable information (consent forms) and ID allocation will be kept separately from the collected data and will be stored on a password protected document. Audio recordings of interviews and focus groups will be password protected and deleted once transcripts have been checked for accuracy. Members of the

research team will be familiar with the requirements of the General Data Protection Regulation (GDPR) and the UK Data Protection Act (DPA) and will comply with these throughout the study.

Access to Final Study Dataset

The team will have exclusive use of the data until key publications from the work described in this proposal have been accepted for publication. This will be within an 18-month period of the study, or 6 months after the study's end date (whichever is earlier). The transcripts from the focus groups and interviews stored in Bristol University's Research Data Repository will be 'restricted', meaning that the Research Data Service will review applications to access these data to check that the applicant is a researcher affiliated to an institution and for what purpose the data are to be used. External users will be bound by the Research Data Access agreement for restricted data: https://drive.google.com/drive/folders/1lF1rt2Dp9IzoxZ603FV7KdObujOVhsCl

Dissemination Policy

Results from the study will be published in open access high impact journals targeting primary and secondary care researchers. These articles will be promoted via Twitter, through departmental newsletters and websites, and through mailing lists, such as those used by the University of Bristol, Society for Academic Primary Care and the NIHR School for Primary Care Research (SPCR). Study outputs will also be presented at national and international primary care and mental health conferences. Blogs written by team members summarising findings will be posted online and promoted via Twitter. The PPI group will advise on how best to disseminate findings to different audiences. All participants will be sent a summary of the findings.

Authorship Eligibility Guidelines

The International Committee of Medical Journal Editors (ICMJE) guidelines for authorship will be followed. They recommend that authorship be based on the following 4 criteria:

- Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND
- Drafting the work or revising it critically for important intellectual content; AND
- Final approval of the version to be published; AND
- Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Bertolote J, McGorry P. Early intervention and recovery for young people with early psychosis: consensus statement. *Br J Psychiatry Suppl* Aug 2005;48:s116-119.

2. Collaborators. GBoDS. Global, regional and national incidence, prevalence and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic review for the Global Burden of Disease Study 2013. *Lancet* 2013;386:743-800.

3. Perala J, Suvisaari J, Saarni SI. Lifetime prevalence of psychotic and bipolar I disorders in a general population. *Arch Gen Psychiat* 2007;64(1):19-28.

4. Rossler W, Salize HJ, van Os J, Riecher-Rossler A. Size of burden of schizophrenia and psychotic disorders. *Eur Neuropsychopharmacol* Aug 2005;15(4):399-409.

5. Kovacs G, Almasi T, Millier A, et al. Direct healthcare cost of schizophrenia - European overview. *Eur Psychiatry* Feb 2018;48:79-92.

6. Addington D, Addington MDJ, Patten S. Relapse rates in an early psychosis treatment service. *Acta Psychiat Scand* Feb 2007;115(2):126-131.

7. Singh SP, Grange T. Measuring pathways to care in first-episode psychosis: A systematic review. *Schizophrenia Research* Jan 1 2006;81(1):75-82.

8. Skeate A, Jackson C, Birchwood M, Jones C. Duration of untreated psychosis and pathways to care in first-episode psychosis - Investigation of help-seeking behaviour in primary care. *Brit J Psychiat* Sep 2002;181:S73-S77.

9. Nazareth I, King M, Haines A, Rangel L, Myers S. Accuracy of Diagnosis of Psychosis on General-Practice Computer-System. *Brit Med J* Jul 3 1993;307(6895):32-34.

10. Kirkbride JB, Errazuriz A, Croudace TJ, Morgan C, Jackson D, Boydell J, Murray RM, Jones PB. Incidence of schizophrenia and other psychoses in England, 1950-2009: a systematic review and meta-analyses. *PLoS One* 2012;7(3):e31660.

11. Sullivan SA, Hamilton W, Tilling K, Redaniel T, Moran P, Lewis G. Early signs and symptoms of psychosis within primary-care: a nested case-control study using electronic primary-care records. *JAMA Network Open* 2018.

12. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, Burroughs H, Jinks C. Saturation in qualitative research: exploring its conceptualization and operationalization. *Qual Quant* 2018;52(4):1893-1907.

13. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res in Psychol* 2006;3(2):77-101.

14. Glaser BG, Strauss A. *The Discovery of Grounded theory: Strategies for Qualitative Research*. Chicago: Aldine Publishing Company; 1967.

15. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. *J Med Internet Res* Jan 30 2015;17(1):e30.

16. Sturges JE, Hanrahan KJ. comparing telephone and face to face qualitative interviewing: a research note. *Qualitative Research* 2004;4:107-118.

17. Taylor AW, Wilson DH, Wakefield M. Differences in health estimates using telephone and door-to-door survey methods--a hypothetical exercise. *Aust N Z J Public Health* Apr 1998;22(2):223-226.

18. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess RG, eds. *Analysing qualitative data*; 1994:173-194.

APPENDICES

Appendix 1 – Required Documentation

1. CVs of research team
2. GP invitation email
3. GP information sheet
4. EIT invitation email
5. EIT clinician information sheet
6. Patient invitation email/letter
7. Patient information sheet and consent form
8. Carer invitation email/letter
9. Carer information sheet and consent form

Appendix 2 – Schedule of Procedures

Procedures	Week
GPs	
Initial invitation email	1
Collection of basic demographic data	1
Confirmation of interview date	2
Verbal informed consent collected	3
Interview conducted	3
EIT clinicians	
Initial invitation email	1
Collection of basic demographic data	1
Confirmation of interview date	2
Verbal informed consent collected	3
Interview conducted	3
Patients	
Practice manager screening	1
Initial invitation email	1
Confirmation of focus group/interview date	2
Verbal informed consent collected	3
Focus group/interview conducted	4
Carers	
Contact details collected from patient participant	1
Initial invitation email	1
Confirmation of focus group/interview date	2
Verbal informed consent collected	3
Focus group/interview conducted	4