Statistical Analysis Plan

The Pharmacogenomics in Depression (PRESIDE) Trial: a double-blind randomised controlled trial of pharmacogenomic-informed prescribing of antidepressants on depression outcomes in patients with major depressive disorder in primary care

1. Administrative details

1.1. Version Number and Date

September 2024 Version 1.0

1.2. Trial registration and ethics approval

Registered prospectively with the Australian and New Zealand Clinical Trial Registry (ACTRN12621000181808) on 22nd February 2021.
https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=380870

Ethical approval was granted by the Human Research Ethics Committee at the University of Melbourne in Melbourne, Australia (Ethics ID: 20626). The Australian Government via Services Australia's External Request Evaluation Committee approved the release of Medicare Benefits Scheme and Pharmaceutical Benefits Scheme data (ID: RMS1620).

1.3. Funding acknowledgement

This work is funded by the National Health and Medical Research Council (NHMRC) (MRF1200060). The design, management, analysis and reporting of the trial are independent of the funders.

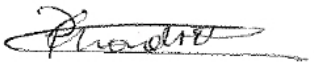
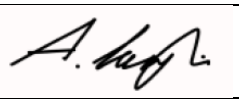
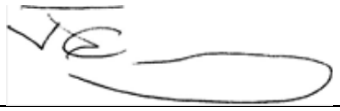
1.4. Acknowledgements

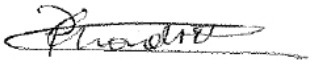
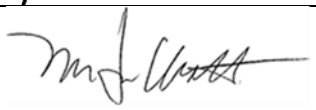
The PRESIDE Trial Investigators and members of the Steering Committee.

1.5. Protocol reference

Saya, S., Chondros, P., Abela, A., Mihalopolous, C., Chatterton, M-L., Gunn, J., et al (2023). The Preside (Pharmacogenomics in Depression) Trial: a double-blind randomised controlled trial of pharmacogenomic-informed prescribing of antidepressants on depression outcomes in patients with major depressive disorder in primary care. *Trials*, 24(1):342[1] https://doi.org/10.1186/s13063-023-07361-6

1.6. Roles and responsibilities – signatory names and contribution

Name	Role	Affiliation	Signature	Date
A/Prof Patty Chondros	Lead Biostatistician (Contact person)	Department of General Practice and Primary Care, University of Melbourne E-mail: p.chondros@unimelb.edu.au		2Oct2024
Adrian Laughlin	Biostatistician	Department of General Practice and Primary Care, University of Melbourne		8Oct2024
Prof Jon Emery	Principal Investigator	Department of General Practice and Primary Care & Centre for Cancer Research, University of Melbourne		02Oct2024

Name	Role	Affiliation	Signature	Date
A/Prof Patty Chondros	Lead Biostatistician (Contact person)	Department of General Practice and Primary Care, University of Melbourne E-mail: p.chondros@unimelb.edu.au		2Oct2024
Dr Sibel Saya	Project manager/Trial co-ordinator	Department of General Practice and Primary Care & Centre for Cancer Research, University of Melbourne		02Oct2024
Dr Mary Lou Chatterton	Health economist	School of Public Health and Preventive Medicine, Monash University		08Oct2024
Prof Cathy Mihalopolous	Lead Health economist	School of Public Health and Preventive Medicine, Monash University		08Oct2024

1.7.Roles and responsibilities – non-signatory names and contribution

Name	Role	Affiliations
Georgia Ramsay	Research Assistant	Department of General Practice and Primary Care & Centre for Cancer Research, University of Melbourne, VIC Australia

1.8.SAP Version History

This is the first version of this SAP approved for publication on 25 September 2024

Protocol version	Updated SAP version no.	Section number changed	Description of and reason for change	Date changed

Contents

1. Administrative details 1
   1.1. Version Number and Date 1
   1.2. Trial registration and ethics approval 1
   1.3. Funding acknowledgement 1
   1.4. Acknowledgements 1
   1.5. Protocol reference 1
   1.6. Roles and responsibilities – signatory names and contribution 1
   1.7. Roles and responsibilities – non-signatory names and contribution 2
   1.8. SAP Version History 2
2. List of Abbreviations 5
3. Synopsis 7
   3.1. Background and rationale 7
   3.2. Trial objectives and hypothesis 8
4. Trial methods 9
   4.1. Trial design 9
   4.2. Eligibility criteria for participants 9
   4.3. Interventions 9
   4.4. Randomisation 10
   4.5. Sample size 10
   4.6. Framework 10
   4.7. Statistical interim analyses and stopping guidance 10
   4.8. Timing of outcome assessments 10
   4.9. Trial protocol modifications 11
5. General Statistical Methodology 12
   5.1. Multiple testing 12
   5.2. Adherence to the intervention 13
   5.3. Protocol Deviations 13
   5.4. Analysis populations 13
   5.5. Handling of missing data 14
   5.6. Assessment of timely completion of self-reported outcome measures 15
   5.7. The Estimand framework 15
6. Trial Population 15
   6.1. Screening Data 15
   6.2. Summary of Eligibility Criteria and Recruitment 15
   6.3. Withdrawal/Follow-up – level of withdrawal 16

6.4. General practice characteristics 16
6.5. Participant characteristics 17
Intercurrent events 18
Measure of impact on COVID-19 pandemic on quality of life at 26 weeks 18
7. Estimands for Primary and secondary objectives 19
7.1. Primary Objective 19
7.2. Secondary Objective 1 20
7.3. Secondary Objective 2 20
7.4. Secondary Objective 3 20
7.5. Secondary Objective 4 21
7.6. Secondary objective 5 21
7.7. Intercurrent events for Secondary Estimands 1 to 5 21
7.8. Outcome for Descriptive objectives 6 22
7.9. Outcome for Descriptive objective 7 22
8. Statistical analysis 22
8.1. Descriptive analysis 22
8.2. General practice characteristics 22
8.3. Participant demographics and outcomes measured at baseline 23
8.4. Analysis of Primary and secondary outcomes 23
9. Health Economics Analysis 27
9.1. Health Economics Objectives 27
9.2. Analysis Framework 27
9.3. Costs 28
9.4. Outcomes 29
9.5. Analysis method 30
9.6. Sensitivity analyses 30
10. Data management and workflow 30
10.1. Timing of final analysis and outcome assessment 31
10.2. Longer term administrative data collection 31
11. Statistical software and technical details 31
12. References 32
13. Appendices 34

2. List of Abbreviations

The following abbreviations are used in this Statistical Analysis Plan (SAP).

Abbreviation	Term
2SLS	Two-stage least squares
AI	Associate Investigator
ATG	Australian Therapeutic Guidelines
ANZCTR	Australian and New Zealand Clinical Trials Registration
AQoL-4D	Assessment of Quality of Life-4D
CACE	Complier Average Causal Effect
CI	Chief Investigator
cLDA	Constrained longitudinal data analysis
CPIC	Clinical Pharmacogenetics Implementation Consortium
CSV	Comma Separated Values
DNA	Deoxyribonucleic Acid
DPWG	Royal Dutch Pharmacogenetics Working Group
FIBSER	Frequency, intensity, burden of side effect rating
GCP	Good Clinical Practice
GLM	generalised linear model
GP	General Practitioner/General Practice
GP EMR	General Practice electronic medical records
HREC	Human Research Ethics Committee
ICERs	Incremental cost-effectiveness ratios
ICH	International Conference on Harmonisation
MAR	Missing At Random
MARS-5	Mediation Adherence report scale
MBS	Medicare Benefits Scheme
MDD	Major Depressive Ddisorder
NHMRC	National Health and Medical Research Council
PAS	Population Analysis Set
PBS	Pharmaceutical Benefits Scheme
PGx	Pharmacogenomics
PHQ-9	Patient Health Questionnaire - 9
PI	Principal Investigator
PRESIDE	Pharmacogenomics In Depression
QA	Quality Assurance
QALY	Quality Adjusted Life Years
QC	Quality Control
RA	Research Assistant
RCT	Randomised Control Trial
REDCap	Research Electronic Data Capture
RGO	Research Governance Office
RUQ	Resource Use Questionnaire
SA	Services Australia
SAP	Statistical Analysis Plan
SoA	Schedule of Assessments
SOP	Standard Operating Procedure
SNRI	Serotonin-Noradrenaline Reuptake Inhibitor
SSRI	Selective Serotonin Reuptake Inhibitor

TGA	Therapeutic Goods Administration
UOM	University of Melbourne
USA	United States of America
VicReN	Victorian Primary Care Practice-based Research and Education Network

3. Synopsis

3.1. Background and rationale

Major depressive disorder (MDD) is a leading cause of non-fatal burden of disease, affecting at least 264 million people worldwide with Australia ranking second in MDD prevalence. Most people with MDD are identified, treated, and followed up by general practitioners (GPs). Therefore, interventions to improve the effectiveness of MDD management have the greatest chance of impact when focused on primary care.

MDD is typically treated with a combination of antidepressant medications such as selective serotonin reuptake inhibitors (SSRIs) or serotonin-noradrenaline reuptake inhibitors (SNRIs) and psychological interventions. The choice of antidepressant and dose is often at the discretion of the patient's GP and may be guided by national guidelines such as the Australian Therapeutic Guidelines. However, lack of response to medication and possible side effects are potential issues with pharmacological treatment of MDD.

Recently, pharmacogenomic-based guidelines have been developed to leverage individual genetic information obtained from pharmacogenomic testing (PGx) to optimise antidepressant selection and dose, and to limit the occurrence of side effects. Recent data has shown that antidepressant treatment recommendations based on genotype-predicted metaboliser phenotypes resulted in greater likelihood of MDD symptom remission compared to usual care. However, these studies were not conducted in a primary care setting, participants were not blinded, and follow-up times were limited to 12 weeks.

Therefore, to address these limitations and gaps in knowledge, a randomised controlled trial was required to determine the effectiveness of PGx-informed antidepressant prescribing recommendations, compared to recommendations based on Australian Therapeutic Guidelines (ATG) for prescribing, in a general practice setting.

This study is a multi-site, individually randomised controlled trial which aims to determine whether the PRESIDE intervention, delivered to Australian general practice patients aged 18-65 years (inclusive) with at least moderate depressive symptoms, results in a change in depression symptom severity between baseline at 12-weeks post-baseline, compared to ATG-informed prescribing.

The PRESIDE intervention incorporates the provision of individually tailored antidepressant reports which provide recommendations to aid in the optimisation of antidepressant treatment of depression, based on PGx. The report is delivered to the patient's GP for consideration by the GP and the patient. GPs of patients in the control arm are sent a similar report containing ATG recommendations.

The PRESIDE trial will also determine the impact of the PRESIDE intervention on symptom remission, symptom response, side effect frequency, medication adherence, patient quality of life, changes in medication and cost effectiveness compared to existing guidelines (control arm).

The trial protocol paper provides further detail on the study rationale and trial design, including the setting, recruitment, eligibility, the intervention procedures, sample size calculations, and randomisation and allocation. This statistical analysis plan (SAP) should be considered as a companion document that provides a more detailed and technical description of the analyses outlined in the protocol. Details of the analyses for all outcomes (clinical and health economics) are included in this SAP, along with planned sensitivity and supplementary analyses. The methods for the process evaluation will be described elsewhere and results reported is a separate publication.

3.2.Trial objectives and hypothesis

3.2.1. Primary Objective

The primary objective of the PRESIDE trial is to evaluate the impact of individually tailored PGx-informed antidepressant prescribing recommendations provided to general practitioners (intervention) on patients' depressive symptom severity at 12-weeks post-randomisation among Australian general practice patients aged 18-65 years (inclusive) with at least moderate depressive symptoms in the past two weeks at baseline, compared to ATG-informed antidepressant prescribing recommendations (control).

3.2.2. Secondary Objectives

The secondary objectives of the PRESIDE trial are to examine the effect of the PRESIDE intervention compared to ATG-informed prescribing among Australian general practice patients aged 18-65 years (inclusive) with at least moderate depressive symptoms in the past two weeks at baseline, on:

1) Change in depressive symptoms at 4-, 8- and 26-weeks post-baseline.
2) Depressive symptom remission at 12-weeks post-baseline.
3) Depressive symptom response at 12-weeks post-baseline.
4) Medication adherence at 4-, 8-, 12-, and 26-weeks post-baseline using administrative dataset.
5) Number of antidepressant medication changes by 26-weeks post-baseline.

3.2.3. Descriptive Objectives

Amongst Australian general practice patients aged 18-65 years (inclusive) with at least moderate depressive symptoms in the past two weeks at baseline taking an AD at each time point, describe by study arm:

6) self-reported antidepressant side effect frequency at 4-, 8-, 12-, and 26-weeks post-baseline.
7) self-reported medication adherence at 4-, 8-, 12-, and 26-weeks post-baseline.

3.2.4. Objective for Process evaluation

Amongst Australian general practice patients aged 18-65 years (inclusive) with at least moderate depressive symptoms in the past two weeks at baseline taking AD at each time point, investigate the difference between the intervention and control arms in the proportion of participants where the GP prescribing is concordant with the medication recommendations in the antidepressant prescribing report. This objective will be addressed as part of the process evaluation which will be described in a separate document.

3.2.5. Health Economics Objective

To determine the cost effectiveness of the PRESIDE intervention compared to ATG-informed antidepressant prescribing between baseline and 26-week follow-up, among Australian general practice patients aged 18-65 years (inclusive) with at least moderate depressive symptoms in the past two weeks at baseline.

4. Trial methods

4.1. Trial design

The trial design is described in detail in the published study protocol[1]. Briefly, the PRESIDE trial is a multi-site, two-arm parallel, double-blinded (researchers and participants) individually randomised controlled superiority trial. The trial will evaluate the impact of the PRESIDE intervention in general practices in Victoria, Australia, which aims to reduce depressive symptom severity at 12-weeks post-baseline among Australian general practice patients aged 18-65 years (inclusive) with at least moderate depressive symptoms at baseline, compared to ATG-informed antidepressant prescribing. Participants are randomised 1:1, stratified by general practice site and baseline antidepressant use, to either the intervention arm or the control arm.

4.2. Eligibility criteria for participants

To be considered **eligible** for participation in the trial, screened patients are required to meet all the following criteria at baseline:

1) Aged 18-65 years (inclusive)
2) Scheduled to have an appointment with a consented GP within 2 days of being approached for participation in the trial
3) Scored a total of at least 10 on the sum of the 9 items of the Patient Health Questionnaire [2](PHQ-9 ≥ 10), indicating at least moderate depressive symptoms in the past two weeks
4) Able to read and understand English
5) Competent to give informed consent.

Screened patients will be considered **ineligible** if they meet any of the following **exclusion criteria** at baseline:

1) Currently taking antipsychotic medication, except if taking quetiapine ≤ 100 mg PRN for sleep with no history of psychosis
2) Pregnant
3) Report they have experienced suicidal thoughts 'nearly every day' in the past two weeks, as per question nine of the PHQ-9
4) Current diagnosis of dementia
5) Current diagnosis of COVID-19
6) Unavailable for study follow-up over the next six months.

4.3. Interventions

Participants allocated to the PRESIDE intervention will have a personalised report sent to their GP which contains recommendations for antidepressant prescribing (medication and dose) that are informed by PGx, with specific focus on two genes: CYP2D6 and CYP2C19.

Participants allocated to the control arm will have a generic report sent to their GP. The report will contain only general recommendations for antidepressant prescribing (medication and dose) that are informed by ATGs.

For the purposes of blinding, the reports delivered to GPs of participants in both arms will be visually identical and all participants will provide a saliva sample for DNA testing. This ensure that there is no way for GPs or participants to determine the arm to which participants have been allocated, reducing the risk of bias.

In both arms, the reports are provided to GPs approximately 2-3 weeks after the blood sample is collected from participants, at which point participants will visit their GP to consider the recommendations provided. All GPs are informed that the report contains recommendations, and that they should employ standard clinical decision-making in reviewing the recommendations, discussing them with the participant and determining treatment course. The decision to adopt any treatment recommendation is at the discretion of the GPs and participants.

Participant recruitment began on 26/05/2021 and concluded on 28/09/2023. The final participant completed all study activities on 09/04/2024.All participants will be followed-up via questionnaire at 4-weeks, 8-weeks, 12-weeks, and 26-weeks. Additional prescription and administrative data will be collected from GP EMR and linked to the Medicare Benefits Schedule (MBS) and Pharmaceutical Benefits Scheme (PBS) data following the conclusion of study activities for all participants.

4.4.Randomisation

Participants were randomly allocated 1:1 to the intervention or control arms. The allocation sequence was computer-generated and stratified by general practice site and current use of antidepressant medication(s), using permuted blocks of random sizes. To ensure allocation concealment, block sizes were not disclosed. Only researchers who randomise participants and generate antidepressant prescribing reports are unblinded during the trial. Further information regarding randomisation, including information about sequence generation, allocation concealment, implementation and blinding are included in the published protocol.

4.5.Sample size

Full details of the initial sample size are provided in the trial protocol. In brief, the sample size for the PRESIDE trial was based on the Target-D and Link-me randomised controlled trials which recruited general practice patients with depressive symptoms. Based on these trials, we calculated that 672 patients would need to be randomised (336 per arm) to detect a between-arm difference of 0.3 standard deviations in PHQ-9 score at 12-weeks from baseline, with 90% power and an alpha level of 0.05 (or 5.0%) and 30% attrition over 12 weeks. Under the conservative assumption that the observed PHQ-9 standard deviation will be 6, this is equivalent to a between-arm difference of 1.8 (0.3*6) in mean PHQ-9 score at 12-weeks from baseline, representing a minimal meaningful reduction in the mean severity of depressive symptoms.

4.6.Framework

The PRESIDE trial's endpoints are testing for superiority of the intervention compared to the control arm.

4.7.Statistical interim analyses and stopping guidance

No formal interim analyses are planned and there are no plans to stop the trial early as the risk of significant adverse effects is low.

4.8.Timing of outcome assessments

Self-reported primary and secondary outcomes will be assessed at baseline, 4-weeks, 8-weeks, 12-weeks, and 26-weeks post-randomisation. Measures used for the health economics outcomes, namely quality of life (AQoL-4D) and the resource use questionnaire (RUQ) will be assessed at baseline, 12 and 26 weeks. The GP EMR for all participants who provided consent will be audited to collect information about consultations related their mental health from date of consent to 26 weeks after the GP had received the AD prescribing report. PBS and MBS data will be requested

initially for 0 to 6 months for participants post-randomisation; If required, remining data from 7 to 12 months post-randomisation may be requested later.

See Table 2 in the trial protocol[1] for the timing of each endpoint measurement.

4.9.Trial protocol modifications

4.9.1. Inclusion of teletrial recruitment and consent

As described in the trial protocol due to COVID-19 and resulting government restrictions, a large proportion of GP appointments were conducted using telehealth instead of in-person. Although in-person recruitment of participants into the trial started in May 2021, to ensure that we were able to invite eligible individuals into the trial a teletrial recruitment and e-consent process was included from August 2021.

4.9.2. Timing of randomisation

A modification to the time individuals were randomised was made to the trial protocol as outlined in published trial protocol. Briefly, initially randomisation occurred at the time of consent. However, in seven of the 65 randomised participants as of 26th August 2021, the saliva samples provided by the participant did not produce adequate results to develop the PGx-based report. Some affected participants declined to provide a second saliva sample and withdrew from the study. Subsequently, the decision was made to randomise participants after confirmation of complete PGx results, to avoid intervention arm participants being prevented from receiving their antidepressant prescribing report.

4.9.3. Sample size modification

Based on the original sample size calculations, we required 470 (235 per arm) individuals responding at 12 weeks. We inflated the sample size required at baseline to 672 to allow for up to 30% attrition in participants between baseline and 12 weeks follow-up. On monitoring attrition of individuals at each follow-up time, attrition was lower at 20% at 12 weeks than we had anticipated. Therefore, after discussion with the steering committee on the 21st September 2023, sample size was revised to 588 participants to be recruited assuming 20% attrition at 12 weeks.

Further, at the steering committee meeting it was agreed to stop recruiting new participants into the PRESIDE trial, after considering the trial progress to date and impact to the trial duration if participant recruitment continued at a slower accrual rate. Reasons for stopping recruitment of new participants was that we were close to the revised target sample size of 588 participants recruited (with approximately 550 recruited participants at the time of the Steering meeting), and the accrual of participants into the trial had slowed to a few participants a week due to changes of our recruitment strategy within practices post-COVID-19. The slower accrual of participants would have considerably extended the timelines for recruitment of participants into the trial (originally projected to be November 2023) by a further 6 to 8 months, substantially increasing the cost and duration of the trial.

4.9.4. Addition of the COVID-19 impact scale.

The PRESIDE trial began during the COVID-19 pandemic. It is possible that the impact of the pandemic on the population's mental health, including prolonged stay-at-home restrictions implemented in Victoria, Australia where the trial was being conducted, could influence the trial. In November 2021, there was evidence emerging of the mental health impact on the population of the pandemic[3]. At this point, the trial steering committee determined it would be important to collect

data on how COVID-19 had impacted the trial participants' mental health, as it was thought that those who were experiencing mental ill health due to mainly to the impact of the pandemic may be treated differently with antidepressants and/or respond differently to antidepressants compared to those with more long-standing or underlying depression.

As of November 2021, no participants had reached the 26-week endpoint of the trial and so the steering committee proposed to add a measure to determine the perceived impact of the pandemic on participants' mental health. The literature was scoped to find a suitable brief scale. The COVID-19 – Impact on Quality of Life (COV19-QoL) scale was determined to be the most suitable but had only been validated to measure this impact retrospectively over a period of one week[4]. It was decided that given the purpose of the scale was only to compare between arms within the PRESIDE trial, and not to generalise more broadly, that the scope of the scale could be expanded to measure the impact over the past 6 months, i.e., since the participants' enrolment in the trial. This would allow an assessment of whether the two arms within the trial perceived their experience of the pandemic in different ways and allow for adjustment, if so.

4.9.5. Exclusion criteria relaxed

Initially patients were excluded from participating in the study if taking any antipsychotic medication (see Eligibility exclusion criteria Section 4.2). However, on 22nd July 2022, after 204 participants had been recruited, this exclusion criterion was relaxed to allow patients taking ≤ 100mg quetiapine for sleep disturbance, with no history of psychosis, to participate in the study as their inclusion was not considered necessary.

An additional exclusion criterion was added on 17th November 2021 to exclude any patients with a current diagnosis of COVID-19 as a precaution to protect anyone involved in handling the saliva samples from the possibility of contracting COVID-19. These patients may be invited into the trial later if they meet eligibility criteria at a subsequent GP appointment.

4.9.6. Secondary objectives for self-reported AD Side Effects and Medication adherence

Self-reported side effects and adherence to AD medications were only measured if the participant reported currently taking these medications. Thus, we will only describe these two outcomes for participants at each time point using AD medication by study arm and overall, but no statistical testing will be conducted for descriptive objectives 6 and 7 above[5].

5. General Statistical Methodology

5.1. Multiple testing

Hypothesis for the primary objective is defined as follows:

Null hypothesis: There is no difference in the mean change of depressive symptoms at 12-weeks post-baseline between the intervention arm (individually tailored PGx-informed, antidepressant prescribing recommendations provided to general practitioners) and the control arm (ATG-informed antidepressant prescribing recommendations), among Australian general practice patients aged 18-65 years (inclusive) with at least moderate depressive symptoms at baseline.

The **alternative hypothesis**: There is a difference in the mean change of depressive symptoms at 12 weeks post-baseline between the intervention and the control arms, among Australian general practice patients aged 18-65 years (inclusive) with at least moderate depressive symptoms at baseline.

Similarly, the following hypotheses are defined for secondary trial outcomes:

Null hypothesis: There is no difference between the intervention and control arms

Alternative hypothesis: There is a difference between the intervention and control arms

All analyses will be two-sided and estimates of the intervention effect will be reported with 95% confidence intervals and p-values. No adjustments for multiplicity of testing to control for overall type I error will be performed.

5.2.Adherence to the intervention

The full intervention includes provision of a saliva sample to study staff for PGx testing, and participant attendance at a follow-up appointment with a GP to discuss the recommendations in their antidepressant prescribing report. Therefore, incomplete adherence will be defined at three levels:

1) Participant does not provide a saliva sample for PGx testing.
2) PGx results are not available for the participant due to sample failure and a repeat sample unable to be collected.
3) The antidepressant prescribing report is not provided to the participant's GP within 3 weeks of baseline PHQ-9 measure (from entry into the trial)

Note, we will be able to determine if the PGx report is delivered GPs, but we will not be able to determine with certainty if the report was seen by the GP. We will however record whether the GP discussed the report with the participant.

The number and proportion of participants meeting each adherence criteria will be reported overall, and by study arm.

5.3.Protocol Deviations

All protocol deviations will be reported, assessed for seriousness, and classified in a blinded review. Deviations that may seriously affect trial quality, effectiveness of the intervention, or participant well-being will be reported.

5.4.Analysis populations

Box 1 provides a description of the population analysis dataset (PAS) to be used for primary and secondary outcomes, and for which analyses they are to be used, as per analyses described in Section 8.

Box 1	Description[1]
Analysis dataset 1 *For analysis of the primary estimand and for the secondary estimands*	PAS 1: The intention to treat (ITT) population, which includes all randomised patients who will be analysed according to the trial arm that they were randomised, and do not withdraw their unprocessed data from the trial[1]. Modified PAS 1: Modified ITT (mITT) population will be defined as above but exclude participants not eligible for the study at the time of randomisation based on the one or more eligibility criteria (i.e. was a randomisation error and would not have been included).
Analysis dataset 2 *For analysis of secondary estimands 2 and 3 if missing responses at 12 weeks is less than 5%; to be determined after blinded review of missing data (See Section 5.5 below)[6]*	PAS 2: All randomised participants with outcome data at 12-week follow-up who do not withdraw their unprocessed data from the trial[1].
Analysis dataset 3 *Additional sensitivity analysis if overall proportion of deaths in sample up to 12 weeks is greater than 5% in sample (See Section 5.5)*	PAS 3: All randomised patients who do not withdraw their unprocessed data from the trial[1] and who do not die prior to 12-week follow-up.
Analysis dataset 4 *Secondary analysis for participants with actionable drug-gene combination*	PAS 4: All randomised patients with an actionable drug-gene combination, who will be analysed according to the trial arm that they were randomised, and do not withdraw their unprocessed data from the trial.

[1]Note: Data provided by participants who later withdrew consent to use their data will be deleted and not included in the primary data analyses, unless the data had been analysed prior to consent being withdrawn.

5.5.Handling of missing data

The methods for handling missing data will be determined after a blinded review of missing data patterns, the reasons for missing data (association between missingness and other variables), and corresponding plausible assumption about the missingness mechanisms[7]. For instance, participant reported outcomes collected using surveys may be missing data due to attrition of participants over time, whereas compared to outcomes derived using administrative data sources where data may be missing when the participant did not provide consent for use of PBS and MBS data. For the primary objective and secondary objective 1, change in depressive symptoms at 4-, 8-, 12-, and 26-weeks from baseline, the primary strategy for handling the missing data will be to use a likelihood-based constrained longitudinal data analysis model (See section 8.4 for details). This approach assumes that probability of missing data in the outcome variable is conditional on the observed measured data included in the models. That is the missing data mechanism is assumed to missing at random (MAR). Where appropriate, methods used to handle missing data such as the inclusion of additional covariates related to missingness of model variables, multiple imputation, or best-worse case analysis will also be considered for the outcomes if missing data is greater than 5%[6]. Pattern-mixture

model to assess the robustness of the missing data assumptions for the primary outcome may be conducted if it is deemed that the data are not recoverable (i.e. able to be estimated without bias based on observed data)[7]. The SAP will be amended accordingly describing the statistical methods used to handle the missing data.

5.6.Assessment of timely completion of self-reported outcome measures

Self-reported primary and secondary outcomes will be assessed at baseline, 4-weeks, 8-weeks, 12-weeks, and 26-weeks post-randomisation (Section 4.8). Participants may not complete self-reported measures in a timely manner, resulting in self-reported outcome measures not being completed within the nominated assessment. Self-reported measures will be considered to be **outside the assessment window** if they are completed more than +/-14 calendar days from the due date for the questionnaire at each timepoint. The exception is if self-reported measures intended for one timepoint are completed in the assessment window for the following timepoint (e.g., the outcome assessment measures sent to a participant at 4-weeks post-randomisation are completed within the assessment window for the 8-week timepoint) then the survey responses may be considered as the outcome measures for the following timepoint. If multiple outcome measures are completed within the same assessment-window for a single timepoint (e.g., the 4-week and 8-week measures are both completed within the assessment-window for the 8-week outcome assessment) then the survey responses closest to the target date will be used.

5.7.The Estimand framework

Sections 7, 8 and 9 use the estimand framework to describe the analytical approaches for the primary and secondary outcome. An estimand provides the precise description of the effect that will be estimated for each trial objective that will address the research question of the trial. Each estimand consists of five elements: target population, intervention, comparator, outcome (variable of interest), the population level summary measure, and the possible events that occur after randomisation (intercurrent events). As part of the estimand framework, we also describe how intercurrent events are handled in the analysis.[8]

6. Trial Population

6.1.Screening Data

Screening data will be recorded electronically in REDCap, a Research Electronic Data Capture web-based software[9]. An electronic recruitment log containing age in years, sex (male, female, other) and PHQ-9 scores and related items of individuals approached for the trial will be kept throughout the recruitment period. Reasons for trial ineligibility or participant refusal (if provided) will also be recorded.

6.2.Summary of Eligibility Criteria and Recruitment

A CONSORT flow diagram (See Appendix) will report the number of individuals who were:

- Screened for eligibility
    - Did not meet inclusion criteria
    - Met exclusion criteria
    - Declined to participate
    - Did not consent/return trial consent
    - Did not consent to data linkage/return consent for data linkage

    - Eligible, but not randomised[1]
- Eligible and randomised

By study arm, after randomly allocated:

- Allocated to study arm at baseline
    - Received allocated intervention
    - Did not receive the allocated intervention[1]
- Follow-up at 4-, 8-, 12-, 26-weeks
    - Discontinued intervention[1]
    - Lost to follow-up[1]
    - Withdrew from further participation
    - Withdrew from completing surveys only
    - Unresponsive (not formally withdrawn)
    - Intercurrent events (Pregnant, died, Antipsychotic medications. etc)[2]
    - Responded to survey at 4-, 8-, 12-, 26-weeks

- Analysed
        - Excluded from analysis[1]

[1]Reasons will be provided.

[2] Note: Some of the intercurrent events may be presented in tables, by study arm and follow-up time.

6.3.Withdrawal/Follow-up – level of withdrawal

Participants can withdraw consent from the trial at any time.

1) Participants can **withdraw their consent to participate in any further trial activities** but maintain consent for any data provided up until the time of withdrawal and may also continue to consent to the collection and use of **the administrative data** (MBS, PBS, GP audit) to be used in the final analysis. Participants who withdraw from further trial participation will receive no further contact regarding the trial or any trial activities.
2) Participants can completely **withdraw consent for the storage and use of any of their unprocessed data, including consent to access to the administrative data.** The data of participants who withdraw consent for the storage and use of their unprocessed data will be deleted as soon as is practicable, except for their age, sex, baseline PHQ-9 score, and study arm status.

The number of participants who withdrew from the study and the type of consent withdrawal will be presented in the CONSORT diagram, overall, by study arm and period of withdrawal. Reasons for withdrawal and loss to follow-up will be presented in the CONSORT diagram, overall, by study arm and follow-up period.

6.4.General practice characteristics

The following data will be collected for each GP clinic site at baseline:

- Index of Relative Socio-Economic Disadvantage (IRSD) based on general practice postcode[10]

- Billing type (bulk billing, private billing or mixed billing)
- Clinic location (Metropolitan/Regional/Rural/Remote) using the Modified Monash Model classification[11]
- Number of GPs in each general practice.

6.5.Participant characteristics

The following data on participant characteristics will be collected at Screening or baseline, prior to randomisation:

Participant characteristics	**Responses**
Postcode of residence	
What is your gender?	Female, Male, Other, Prefer not say
Age at consent (calculated using date of birth)	years
Which language do you mainly speak at home?	English, Arabic, Cantonese, German, Greek, Italian, Mandarin, Spanish, Vietnamese, Other (please specify)
Which ethnicity do you identify most with? (Tick all that apply)	Aboriginal and/or Torres Strait Islander; Central/South Asian; East Asian; European; Near Eastern; Oceanian; Sub-Saharan African; Latin American; Other (please specify); Prefer not to answer
What is the highest level of education you have completed?	Below Year 10; Year 10; Year 11; Year 12 or equivalent; Certificate III/IV; Advanced Diploma / Diploma; Bachelor's Degree; Graduate Diploma / Graduate Certificate; Postgraduate degree
In terms of employment, in a usual week are you..	Working for an employer for wages or salary; Working in your own business for profit or pay; Working without pay in a family business or on a farm; Unemployed, looking for and available to start work; Unemployed, unable to work; None of the above
If answer "Unemployed, unable to work" or "None of the above" then the respondents are asked to complete the following: In a usual week, which of the following best describes your main activity?	Retired or voluntarily; Home duties; Caring for children; Studying; Caring for an ill or disabled person; Working in an unpaid voluntary job; Unable to work due to own illness, injury, or disability; Other
Do you live alone?	Yes, No
If No above: Who do you usually live with? (Tick all that apply)	Husband or wife; Defacto partner; My child/ren; My partner's child/ren; My parent/s; Unrelated flatmate or co-tenant; Other relationship; Other
About your health	
Are you taking any medications for your mental health?	Yes, No
If yes, which medications? (Tick all that apply)	See REDCap data dictionary for list
Are you taking medications for other health problems? This includes prescribed and non-prescribed medications.	Yes, No
If yes, What medications are you currently taking (please list all)	Text
Have you taken antidepressant medication in the past?	Yes, No
If yes, what antidepressant medication(s) have you taken in the past? (Tick all that apply)	See REDCap data dictionary for list
Are you a current smoker?	Yes, No

How often do you have a drink containing alcohol?	Never; Monthly or less; 2-4 times a month; 2-3 times a week; 4 or more times a week
How many standard drinks do you have on a typical day when you are drinking?	1 or 2; 3 or 4; 5 or 6; 7 or 9; 10 or more
Are you a current cannabis user?	Yes, No
Are you a current medicinal cannabis user?	Yes, No

Individuals' Index of Relative Socioeconomic Disadvantage (IRSD) will be derived using postcode of residence[10].

Other measures collected at screening and baseline include:

- Mode of participant recruitment (in person or teletrial)
- Mode consent obtained (in person and eConsent)
- DNA swab completed (yes, no, pending)
- DNA collection mode (In person or via Zoom)
- Measures of the outcomes described in Section 7 below
- Health care service utilisation, including non-pharmacological therapies collected using the resource use questionnaire (RUQ). Additional data of health care service used will be captured via administrative dataset (GP Audit, PBS and MBS)

Full details of data collected at screening and baseline via self-report are provided in the REDCap data dictionaries for the PRESIDE trial. Data dictionaries are available on request.

Intercurrent events

Following events that occur after randomisation and within the trial duration will also be captured:

1) Participant death and time point, when reported in the GP records
2) Women became pregnant, when reported in the GP records
3) Initiation of concomitant non-pharmacological therapies and time initiated
4) Initiation of concomitant antipsychotic treatment and time initiated

Measure of impact on COVID-19 pandemic on quality of life at 26 weeks

COVID-19 pandemic in quality of life (COV19-QoL) scale measured at 26-week follow-up[4]. The COV19-QoL scale was developed to capture the effect of COVID-19 on individuals' quality of life in relation to mental health over the past 7 days. There are six items as shown below, with responses on a five-point Likert scale (1=completely disagree to 5=completely agree). For the purposes of the trial, we modified the period from 7 days to 26 weeks, the duration of the trial. Total score is the average of responses to the 6 items, and ranges between 1 to 5, where higher scores indicate greater perceived impact of the pandemic on participants quality of life.

COV19-QoL	**Responses**
We are interested in the impact that the COVOD-19 pandemic has had on your mental-health. Please choose the number that best represents the degree of your agreement with the statements provided below. Please keep in mind that your estimates reflect your feelings and thoughts during the past 6 months due to the spread of the coronavirus (COVID-19).	
1) I think my quality of life in lower than before	1=Completely disagree
2) I think me mental health has deteriorated	2=Disagree
3) I think my physical health may deteriorate	3=Neither agree nor disagree
4) I feel more tense than before	4=Agree

5) I feel more depressed than before 6) I feel that my personal safety is at risk	5=Completely agree

7. Estimands for Primary and secondary objectives

This section describes the primary and secondary estimands, including population-level summary measure, the definition and derivation of the outcome of interest for each objective, potential intercurrent events and how they will be handled in the analysis. The analytical methods for the primary and secondary outcomes to estimate the are provided in Section 8 and 9.

7.1.Primary Objective

Primary estimand: The difference between intervention and control arms in the mean change in depressive symptoms at 12weeks compared at baseline, measured using the Patient Health Questionnaire (PHQ-9)[12], among Australian general practice patients aged 18-65 years (inclusive) with at least moderate depressive symptoms (PHQ-9 score ≥ 10) at the time of recruitment.

The **PHQ-9 score** consists of nine items, each measured on a four-point Likert scale, that assess various aspect of patient health over the preceding two-week period. Responses from each item range numerically from 0 to 3, with a higher score indicating poorer health. The depressive symptom scores are the aggregate of the responses for all nine items, which range from 0 to 27, with higher scores indicating more depressive symptoms. If two or fewer items are missing a response, then missing values will be replaced with the mean of the completed items[2]. If responses are missing for more than two items, then no PHQ-9 score will be calculated.

The **outcome** is the mean change in PHQ-9 score between baseline and 12-weeks, post-baseline. For each individual, change in PHQ-9 score will be calculated as the difference between the PHQ-9 score at 12 weeks and PHQ-9 score at baseline. Negative values of the change in PHQ-9 scores at 12 weeks indicate decrease in severity of depressive symptoms in the past two weeks at 12 weeks compared to baseline.

Intercurrent Events for the Primary Estimand

Death: During the trial period, participants may die prior to 12-week follow-up, precluding the existence of primary outcome data. We will adopt a **hypothetical approach** using statistical methods to estimate the parameter of interest as if 12-week outcome data had been observed for participants who experience death prior to 12-week follow-up. A **hypothetical strategy** is considered reasonable for handling intercurrent events when considering patient-reported outcomes if the number of deaths is expected to be low and death is not a result of the intervention[13] (ref).

Pregnancy: Participants may become pregnant during the trial period, prior to the 12-week follow-up, affecting interpretation of the primary outcome. Pregnancy may directly or indirectly affect the interpretation of the outcome as it may affect depressive symptoms due to biological changes. However, considering pregnancy from a pragmatic perspective, we will adopt a **treatment policy** approach.

Non-pharmacological therapies: Initiation of concomitant non-pharmacological therapies are expected throughout the duration of the trial and may affect interpretation of the primary outcome. Patients who begin, for example, psychological therapy may experience changes to their depressive symptoms independent of the intervention. However, non-pharmacological interventions are a

regular form of care for individuals with depressive symptoms, and many study participants may be under the care of a psychologist/psychiatrist upon entering the study (and thus, these effects will be distributed approximately equally between arms due to random allocation). Therefore, patients who begin concomitant non-pharmacological therapies during the trial will be handled using a **treatment policy** approach, where non-pharmacological therapies are considered part of the intervention to be evaluated and becomes part of the definition of the trial treatment.

Antipsychotic medications: Initiation of concomitant antipsychotic treatment may also occur throughout the duration of the trial, affecting the interpretation of the outcome. Such medications are often used to treat individuals with more severe mental ill-health, including conditions such as schizophrenia. Such medications may impact the measured outcomes. As above, the patients who are prescribed a**ntipsychotic medications** therapies during the trial will be handled using a **treatment policy** approach.

7.2.Secondary Objective 1

Secondary Estimand 1: The difference between intervention and control arms in the mean change in depressive symptom score from baseline, measured using the PHQ-9, at (i) 4-weeks, (ii) 8-weeks, and (iii) 26-weeks, among Australian general practice patients aged 18-65 years (inclusive) with at least moderate depressive symptoms (PHQ-9 score ≥ 10) at baseline.

The **outcome** for Secondary Estimand 1 is the mean change in PHQ-9 score at (i) 4-weeks, (ii) 8-weeks, and (iii) 26-weeks from baseline PHQ-9 score, respectively.

Similar to the primary outcome, change in PHQ-9 score at each time point will be calculated as the difference between the PHQ-9 score at each time point minus the PHQ-9 score at baseline.

7.3.Secondary Objective 2

Secondary Estimand 2: Between the intervention and control arms the (i) difference in the proportion of participants (absolute measure), and (ii) the odds ratio of participants (relative measure), with depressive symptom remission (defined as PHQ-9 < 5) at 12-weeks post-baseline among Australian general practice patients aged 18-65 years (inclusive) with at least moderate depressive symptoms (PHQ-9 score ≥ 10) at baseline.

The **outcome** is proportion of participants with remission of depressive symptoms at 12-weeks post-baseline, defined as a PHQ-9 score at 12 weeks of less than 5.

A binary indicator variable will be generated containing the value 0 for participants who did not experience depressive symptom remission and a value of 1 for those who did experience depressive symptom remission. If PHQ-9 score at 12 weeks is missing the indicator variable for remission will be coded as missing.

7.4.Secondary Objective 3

Secondary Estimand 3: Between the intervention and control arms in the (i) difference in the proportion of participants (absolute measure), and (ii) the odds ratio of participants (relative measure), with depressive symptom response (defined as ≥50% reduction in baseline PHQ-9) at 12-weeks post-baseline among Australian general practice patients aged 18-65 years (inclusive) with at least moderate depressive symptoms (PHQ-9 score ≥ 10) at baseline.

The **outcome** is proportion of participants with a depressive symptom response at 12-weeks post-baseline, defined as a minimum 50% reduction in baseline PHQ-9 at 12 weeks.

Percent reduction in baseline PHQ-9 score at 12 weeks will be calculated by dividing the change in PHQ-9 score from baseline at 12 weeks with the baseline PHQ-9 score, and then multiplied by 100. A binary indicator variable will be created where 0 will indicate the participant did not experience a depressive symptom response ($percent\ reduction < 50\%$) and 1 will indicate the participant did experience a depressive symptom response ($percent\ reduction \geq 50\%$). If PHQ-9 score at baseline or 12 weeks is missing the indicator variable will be coded as missing.

7.5.Secondary Objective 4

Secondary Estimand 4: Difference (absolute measure) and ratio (relative measure) of the antidepressant medication possession ratio (MPR)[14] between intervention and control, among Australian general practice patients aged 18-65 years (inclusive) with at least moderate depressive symptoms (PHQ-9 score ≥ 10) at baseline.

Th **outcome MPR** is calculated as the sum of days' supply of any AD medication dispensed (and assumed possessed) divided by the total days the participant was observed during the trial period, over the observed period from baseline to 26-weeks post-baseline, as determined from prescriptions recorded in the GP EMR and/or in the PBS data. The MPR will not be calculated for individual medications but will be a combined medication possession ratio for any AD medication during the trial period. Thus, given n is the total number of trial participants, the $MPR_i$for the i^{th} individual, where $i = 1, 2, 3, \dots, n$ will be defined as:

$$MPR_i = \frac{\sum_{k=1}^{t_i} x_{i_k}}{t_i}$$

where x_{i_k} is a binary indicator variable coded as 1 indicating possession of the AD medication on the k^{th} day or 0 otherwise for i^{th} participant; t_i is the total number of days the i^{th} participant is observed, with a maximum number days of 182 days (26 weeks) from baseline.

7.6.Secondary objective 5

Secondary estimand 5: Difference (absolute measure) and ratio (relative measure) of the rate of antidepressant medication changes in the intervention arm compared to the control arm, among Australian general practice patients aged 18-65 years (inclusive) with at least moderate depressive symptoms (PHQ-9 score ≥ 10) at baseline.

The **outcome** is the rate of antidepressant medication (AD) changes from baseline to 26-weeks post-baseline. It is calculated as the count of AD changes, as recorded within the GP EMR and/or linked PBS data, divided by the total time participants are observed during the trial period. An AD medication change is defined as a recorded prescription for an AD medication that does not match AD medication at the last recorded prescription. Neither initiation nor cessation will be considered as instances of AD medication changes.

7.7.Intercurrent events for Secondary Estimands 1 to 5

1) **Death**: As for the primary estimand, death during the trial period may prevent observation of outcome data. For all estimands utilising patient reported outcome measures, death will be handled using a **hypothetical approach**.
2) **Pregnancy**: As with the Primary Estimand, all Secondary Estimands will handle pregnancy using a **treatment policy** approach due to the pragmatic nature of the trial.
3) **Non-pharmacological therapies**: As with the Primary Estimand, for any concomitant non-pharmacological therapies, a **treatment policy** approach will be taken for all Secondary Estimands. This acknowledges that treatment for mental ill-health involves a variety of

interventions, such as psychological therapy, which are commonly used in conjunction with pharmacological therapies.

4) **Antipsychotic medications:** As for non-pharmacological therapies, initiation of antipsychotic medications will be handled using a **treatment policy**.

7.8.Outcome for Descriptive objectives 6

Outcome is self-reported antidepressant side effect score using the Self-Rated Global Measure of the Frequency, Intensity, and Burden of Side Effects Rating (FIBSER)[15] at (i) 4-weeks, (ii) 8-weeks, (iii) 12-weeks, and (iv) 26-weeks post-baseline. The three items are only asked for participants who report currently taking prescribed AD medications for depression.

The **FIBSER scale** consists of three items, each measure one element of antidepressant medication side effects: frequency, intensity, and burden/impairment of taking AD medication in the last week. Each item is measured on a seven-point Likert scale with numerical scores ranging from 0 to 6. Higher scores indicate that the antidepressant medication is having a greater adverse effect for that element.

7.9.Outcome for Descriptive objective 7

Outcome is self-reported adherence to prescribed antidepressants using the Medication Adherence Report Scale (MARS-5)[16] at (i) 4-weeks, (ii) 8-weeks, (iii) 12-weeks, and (iv) 26-weeks post-baseline. These items are only asked for participants who report currently taking prescribed AD medications for depression.

The MARS-5 instrument assesses self-reported adherence to medication using five items, each measured on a five-point Likert scale, scored from 1 to 5, with higher scores indicating greater adherence. Each item relates to various aspects of adherence (see below). Adherence is measured on a continuous scale using the MARS-5 score calculated by summing the responses to the five items, which ranges from 5 to 25, where higher scores indicate greater adherence to antidepressants.

MARS-5	**Responses**
Many people find a way of using their medicines which suits them. This may differ from the instructions on the label or from what the doctor had said. Here are some ways in which people have said they use their medicines. For each statement, please tick the box which best applies to you.	
1) I forgot to take the medicine	5=never
2) I alter the dose of the medicine	4=rarely
3) I stop taking the medicine for a while	3=sometimes
4) I decided to miss a dose	2=often
5) I take less than instructed	1=very often

8. Statistical analysis

8.1.Descriptive analysis

8.2. General practice characteristics

GP characteristics will be summarised using descriptive statistics.

8.3.Participant demographics and outcomes measured at baseline

Participant demographics and baseline measures of the outcomes will be summarised using descriptive statistics, overall and by study arm to identify any potential imbalance between arms at baseline. The mean and standard deviation will be used to summarise continuous variables (e.g., age), or median and interquartile range (IQR; 25th and 75th percentile) if the variable has a skewed distribution. Counts and percentages will be presented for categorical variables.

Descriptive statistics will also be used to describe, as appropriate:

1) The distribution of age (years) and sex of trial participants versus non-consenting screened patients
2) Baseline characteristics, such as PHQ-9 score, of participants who remained in the study versus those who withdrew or were lost to follow-up
3) Baseline characteristics, such as PHQ-9 score, of participants with and without missing primary endpoint data
4) Baseline characteristics, such as PHQ-9 score, of participants who did and did not respond to surveys at 4-, 8-, and 26-week follow-up.
5) Intercurrent events by study arm and overall
6) COV19-QoL impact scale by study arm and overall

8.4.Analysis of Primary and secondary outcomes

Details of all statistical analyses relating to the primary and secondary objectives, including the primary, sensitivity and supplementary analyses, are described in this section. Analyses related to the Health Economics outcomes are described in Section 9.

8.4.1. General principles

In all regression analyses, study arm, depressive symptom severity and randomisation stratification factors (GP site and current antidepressant use) will be included as covariates. The primary analyses for the primary and all secondary outcomes will use PAS 1 (Section 5.4), an intention-to-treat analysis population, where all randomised participants are included in the analysis according to their allocated study arm, irrespective of the level of adherence to the intervention. Exception is when if consent for use of any data collected is withdrawn prior to the analysis. The primary analysis all self-reported responses will be included in the analysis, even if they are received outside of the corresponding assessment window as outlined in Section 5.6.

8.4.1.1. Primary analysis

For the **Primary Estimand** and **Secondary Estimand 1**, a constrained longitudinal data analysis (cLDA) with the response variable consisting of all PHQ-9 scores measured at each time (Baseline, 4-, 8-, 12-, and 26- weeks) will be used to estimate the difference in the mean change in depressive symptom scores between the intervention and control arms, from baseline to each of the follow-up time (namely, 4-weeks, 8-weeks, 12-weeks and 26-weeks). A linear mixed-effects model using the restricted maximum likelihood with random effects for individuals to account for the correlation between repeated measures on the same individual, and fixed effects for study arm (1=intervention, 0=control) and follow-up time (baseline, 4, 8, 12 and 26 weeks), and an interaction term for study arm and time. The model will adjust for depressive symptoms at baseline, where the estimates for the mean depressive symptoms at baseline will be constrained to be equal between the two study arms. The variance-covariance structure for the repeated measures within individuals will be defined as unstructured. Alternative structures will be considered there is non-convergence. Under this

model, the inference for the missing data mechanism is MAR. The estimated between-arm difference in change in mean depressive symptom score at 4-, 8-, 12- and 26-weeks will be presented with the 95% confidence interval and p-value.

Secondary Estimands 2 and 3 with a binary outcome, the difference in proportions (absolute measure) will be estimated using a generalised linear model (GLM) with the identity link function and binomial distribution. The odds ratio (relative measure) will be estimated using logistic regression. Both models will adjust for depressive symptoms measured at baseline and randomisation stratification variables (GP site and current antidepressant use). If the model used to estimate the absolute difference between study arm fails to converge, the difference in proportions will be derived from the logistic regression model[17,18].

The absolute (difference in proportions between the intervention and control arms) and relative (odds ratio) estimated intervention effects for the primary outcome will be reported with 95% confidence intervals and p-values. P-values for the binary outcomes will be estimated using the logistic model.

Secondary Estimands 4 and 5 with a count outcome

The difference in the outcomes (namely, (i) MPR, and (ii) rate of AD changes) between the intervention and control arms and the ratio will be estimated using Poisson mixed effects model (or a Negative Binomial mixed effects model if overdispersion is detected), with fixed effects for study arm, baseline depressive symptoms and randomisation stratification factors (GP site and current antidepressant use).

The estimated intervention effect will be reported as differences in (i) MPR, and (ii) rate of AD changes between the intervention and control arms and the respective ratios with 95% confidence intervals and p-values.

8.4.2. Sensitivity analyses

Sensitivity Analysis 1: Adjustment for additional covariates

In addition to the variables included in the models for the primary analyses, each outcome will be analysed adjusting for additional covariates. The following variables will be specified as fixed effects in the models for the primary and secondary outcomes:

Adjusted model 1: age and sex
Adjusted model 2: age, sex plus ancestry defined by self-reported ethnicity, if ethnicity at baseline is imbalanced between study arms. As CYP2C19 and CYP2D6 phenotype frequencies can vary by ancestry. If unexpected imbalance of other baseline factors between arms identified after a blinded review may be included also be included as fixed effects in the adjusted model.
Adjusted model 3: include as a potential confounder the COV19-QoL scale measuring COVID-19 pandemic in quality of life if after a blinded review is found to be imbalanced between the two study arms.

Sensitivity analysis 2: Self-reported outcomes measured outside assessment-window

Repeat primary analysis for the **Primary estimand and Secondary Estimands 1 to 3** where survey outcome responses are outside the assessment window (as described on Section 5.6) are set to missing. Method used to handle the missing data are described in Section 5.5.

Sensitivity analysis 3: Hypothetical Approach for intercurrent event related to the initiation of antipsychotic medications for Primary estimand and Secondary Estimands 2, 3 and 4

Individuals currently using of antipsychotic medication were ineligible to participation in the trial (as described in Section 4.2). In this sensitivity analysis we will use the **hypothetical approach** for this intercurrent event, to evaluate the impact on estimating the effect of the intervention had participants not initiated antipsychotic treatment. For participants where **antipsychotic medications** are initiated within the trial period, we will re-code outcome responses to missing from the time the a**ntipsychotic medications are initiated**. Section 5.5 describes how missing data will be handled in the analysis.

Sensitivity analysis 4: Modified ITT analysis

If appropriate, secondary estimands will be considered which will be the same as the primary and secondary estimands (Section 7 and Section 8.4.1.1), except with the mITT population, which will exclude participants that did not meet the eligibility criteria at the time of randomisation.

8.4.3. Additional Estimands

Adherence-adjusted analysis

We will perform a supplementary analysis to estimate between-arm intervention effect for primary estimand (primary outcome) and the Secondary estimands 1 to 3 for when the intervention was implemented as intended as described in the trial protocol. **Adherence to the intervention** according to the protocol will be defined at three levels:

1) Participant provided a saliva sample for PGx testing

2) Participant provided a saliva sample for PGx testing

   AND

   Initial saliva sample (or repeat sample) was successfully used to generate PGx results

3) Participant provided a saliva sample
   AND
   Initial saliva sample (or repeat sample) was successfully used to generate PGx results
   AND
   The antidepressant prescribing report is provided to the participant's GP within 3 weeks of baseline PHQ-9 measure (from entry into the trial).

It is expected that the effect of the intervention will be less if saliva samples are not provided by participants, a PGx report is not able to be produced from the saliva sample and no repeat sampling is possible, or if the PGx report is not delivered to the GP.

The intercurrent events of non-adherence across these three levels will be handled using a **hypothetical strategy.** If appropriate, an adherence-adjusted analysis will be carried out using complier average causal effect (CACE) analysis[19-21] for the primary outcome and secondary outcome 2, 3 and 4. A two-stage least squares (2SLS) instrumental variable regression will be undertaken using generated binary indicator variables for adherence, according to each of the definitions provided above, and study arm as the instrumental variable for adherence to the intervention. The model will include the randomisation stratification variables (general practice site and

antidepressant use at baseline). To test the robustness of assumptions underlying the described method, other methods for assessing the impact of incomplete adherence may be utilised[21].

Actionable drug-gene combination

The objective is to explore the effect of the intervention on the primary and secondary outcomes amongst the eligible trial participants who have an actionable drug-gene combination (as defined in Table 1 of protocol paper). We will repeat analysis for primary and secondary estimands, but the target population will be eligible participants with an actionable drug-gene interaction (PAS 4, Section 5.4). We expect that the intervention effect of the primary analysis will be diluted if there is a large proportion of participants without an actionable drug–gene combination. In addition, we will examine the differences in percentage of participants with AD discontinuation between the study arms among those with an actionable phenotype and taking an AD at baseline using the analysis described above for the binary outcome.

8.4.4. Sub-group Analyses

Two exploratory sub-group analyses will be conducted using the same methods as for the primary analyses for change in depressive symptoms scores from baseline at each follow-up each time point (4, 8, 12 and 26 weeks). Each sub-group analysis will include an additional explanatory variable as a fixed effect, modelled with a term for interaction with study arm at each follow-up time point to determine if the effect of the intervention is modified by sub-group. The two additional variables to be included in each model are described below:

1. Prior antidepressant experience: at baseline participants can be categorised into three groups based on previous antidepressant treatment. A participant may have never taken antidepressants at baseline, may have taken antidepressant previously but not taking antidepressants at baseline, or may be currently taking antidepressants at baseline. It is possible that treatment naïve participants may respond differently to treatment compared to those who have previously taken antidepressants or are currently taking antidepressants.
2. Depressive symptom severity at baseline: trial eligibility requires participants to have at least moderate depressive symptoms at baseline (PHQ-9 score≥ 10). However, depressive symptom severity may be further categorised into three categories : moderate depressive symptoms (10-14), moderately severe depressive symptoms (15-19) and severe depressive symptoms (20-27). Participants with varying degrees of severity may respond differently to treatment.

When the p-value for the interaction term between the sub-group variable and the study arm at each time point is less than 0.1 (at one or more of the assessment times 4,- 8-, 12- and 26-weeks), we will present summary statistics and the estimate of the difference in mean change score between the intervention and control arm for each level of the sub-group with 95% confidence intervals and the corresponding p-value for the interaction term between the study arm and the sub-group variable at each follow-up time point. Estimates may also be displayed using forest plots.

If appropriate, the sub-group analysis may also be conducted for the other secondary outcomes.

8.4.5. Descriptive objectives

There are two descriptive objectives for this trial. The outcomes for Objectives 8 and 7 below were only measured for participants who reported taking AD at each follow-up time point

Descriptive objective 6

Count and percentages for each element of antidepressant medication side effects: frequency, intensity, and burden/impairment in the past week using the FIBSER measure at each follow-up-time point for participants taking AD, by study arm and overall.

Descriptive objective 7

Means and standard deviation of adherence to prescribed antidepressants using the MARS-5 score at each follow-up time point for participants taking AD by study arm and overall. If distribution is skewed, we may present the median and interquartile range (IQR; 25th and 75th percentile), minimum and maximum values. The summary statistics may be presented graphically (e.g. boxplots) over time.

Exploratory analysis

Summary statistics for the side effects and adherence measures may be presented separately for each type of AD taken (e.g. Agomelatine, Citalopram, Duloxetine etc). Further we may conduct an analysis to explore whether the severity of the antidepressant medication side effects (frequency, intensity and burden) impact depressive symptom over the trial duration[22].

9. Health Economics Analysis

9.1.Health Economics Objectives

Among Australian general practice patients aged 18-65 years (inclusive) with at least moderate depressive symptoms (PHQ-9 score ≥ 10) at baseline, the difference between intervention and control arms in:

1) mean quality-adjusted life-years (QALYs, calculated using the AQoL-4D utility values and the area under the curve method).
2) Mean total health sector costs (cost of intervention delivery and participant health care service use) from baseline to (i) 12-weeks and (ii) 26-weeks post-baseline.
3) Mean total partial societal costs (cost of lost productivity and health sector costs) from baseline to (i) 12-weeks and (ii) 26-weeks post-baseline.
4) Mean total health sector costs from baseline to 12 and 26 weeks calculated using participant linked MBS and PBS data (sensitivity analysis).

9.2. Analysis F**ramework**

The overall framework for the analysis will be a full economic evaluation using a within-trial method. A health sector perspective will be adopted as the primary perspective and will include costs borne by the government as a third-party payer in addition to out-of-pocket costs incurred by participants when accessing medical care. A partial societal perspective, which includes absenteeism and presenteeism effects on productivity for study participants, will be undertaken as a secondary analysis.

The reference year for the cost analyses will be year 2022/2023. Since the time horizon for the analyses is less than one year, costs and benefits will not be discounted.

9.3.Costs

The health care sector costs include the cost of delivering the individually tailored PGx-informed antidepressant prescribing recommendations, as well as other healthcare services used by study participants from both intervention and control arms of the trial. A partial societal perspective incorporates the cost of productivity loss (including absenteeism and presenteeism). Data will be collected at baseline, 12-weeks and 26-weeks post randomisation on health care service use, out of pocket expenses, time off paid and unpaid work, and estimated impairment while working.

The cost of the PGx-informed prescribing will include the cost of GP training at each clinic, in addition to the cost of sample collection, testing, and provision of results. As all participants provide a saliva sample using the ORAcollect®-DNA OCR100 kit, the cost of collection and PGx testing will be included for both randomised arms in the first base-case analysis. The rationale for including the GP training and testing cost for control arm participants is to provide transparency and completeness, adhering to the methods of the trial. The PGx testing of control arm participants may have led to changes in anti-depressant therapy due to increased awareness. The cost of the collection and testing will be based on the price listed on the Melbourne Pathology website ($197)[23]. A second base-case analysis will exclude the cost of the PGx test for the control arm since the test is a research driven cost for the purpose of maintaining blinding.

Additional health care services that participants, within both the intervention and control arms, access over the course of the trial period for the purpose of managing their mental health will be captured with the self-reported resource use questionnaire (RUQ) and additional individual Medicare Benefits Schedule (MBS) and Pharmaceutical Benefits Scheme (PBS) data. The self-reported resource use will be used in the two base-case analyses.

Health professional visits will be costed using a weighted average cost paid by the government for the corresponding health professional, derived from the Medicare Benefits Schedule (MBS) item reports[24]. Since a standard co-payment for health professional visits is not in place under the MBS, participants were asked to report estimated out of pocket costs paid for these services.
Use of other resources such as books, online therapy or other digital interventions (i.e. apps) and helplines were also reported in the RUQ. The reported out of pocket costs paid by the participants for these services will be included in out-of-pocket costs.

Pharmaceutical Benefits Scheme (PBS) approved ex-manufacturer item prices (1 June 2023) will be used to calculate the government and patient out of pocket costs for mental health medications reported in the RUQ[25].

The reported number of times ambulance services were used by participants will be multiplied by an average ambulance service cost. Any out-of-pocket costs for ambulance services reported by participants will be incorporated into the out-of-pocket category costs.

Emergency department visits will be costed using the Independent Health and Aged Care Pricing Authority (IHACPA) National Weighted Activity Unit (NWAU) calculator for emergency department care multiplied by the 2022/2023 National Efficient Price of $5,797 per NWAU (IHACPA). The out-of-pocket cost for emergency department services was reported by participants and will be added into the total out of pocket cost category.

Hospital stays will be costed using the IHACPA NWAU calculator for acute admitted stays multiplied by the 2022/2023 National Efficient Price of $5,797 per NWAU (IHACPA). The specific diagnostic related groups (AR-DRGs) will be selected based on the reported reason for hospitalisation and the length of stay.

The partial societal perspective incorporates effects on productivity. Participants were asked about the number of hours or days they have taken off from paid and unpaid work. They were also asked to report the number of hours or days when they were bothered by mental health problems while at work along with a question regarding their average capacity during these periods. The human capital approach will be used to value lost paid productivity using an average hourly wage rate calculated from the average weekly earnings reported by the Australian Bureau of Statistics (May 2023) [26]plus 25% overhead costs. Time off from unpaid activities (i.e., household activities) will be valued at 25% of the average wage rate plus overhead costs to represent the value of participants' lost leisure time[27].

Presenteeism will be calculated by multiplying the number of hours reported working but bothered by depression by 10 minus the numeric response regarding the amount of normal work capacity achieved on these days divided by 10. The result will provide the number of hours due to presenteeism which will then by valued using the average wage rate plus overhead costs.

In cases where participants reported days absent or days bothered by depression (rather than hours as this was an option to aid participant recall), an assumption regarding the number of hours worked per day will be used. This will assume a 7.6 hour workday for those working full-time and a 3.8 hour work day for those working part-time.

Following valuation, costs will be aggregated at the following group levels – intervention delivery, health professional consultations, medications, out of pocket costs, acute care services (ambulance, emergency department and hospitalisations), and productivity loss.

9.4.Outcomes

In Australia the preferred outcome measure in health economic evaluations is the quality adjusted life year (QALY) because cost-effectiveness ratios using QALYs have inherent value-for-money connotations. The Australian value set for the Assessment of Quality of Life-4D (AQoL-4D) will be used to derive utility values at baseline, 12- and 26- week time points. The utility values at each time point will be used to calculate total QALYs for each participant using the area under the curve method.

The health economics outcome includes the health-related quality of life (AQoL-4D utility) score at (i) 12-weeks, and (ii) 26-weeks post-baseline. The AQoL-4D is a 12-item instrument consisting of four dimensions (independent living, mental health, relationships, and senses), each containing three items that are scored on a four-point Likert scale, from 1 to 4. A broad description of the method used to calculate the utility score is outlined here. For a detailed explanation, the reader is directed to the AQoL website[28]. The utility score is derived by first generating a 'disvalue' for each response for each item. These disvalues differ by response and by item. An overall disvalue for each dimension is then calculated by applying a series of weights for each item and a weight for each dimension. The final score for each dimension is a continuous (non-integer) value that ranges from 0 to 1, where higher scores are indicative of a better health state. An algorithm is then applied the dimension scores, with a weight for each dimension, that generates a final utility score which ranges from –0.04

indicating a state worse than death (a score of 0 indicates a person is dead) to 1.00 with higher scores indicating higher health-related quality of life. The scoring algorithm for the AQoL-4D utility score allows for one missing value per dimension, and this value is imputed using the mean of the remaining two items for the dimension. If more than two items are missing on a single dimension then the dimension score is not calculated, and subsequently, no utility score is calculated.

The PHQ-9 depression symptoms score will be utilised as an additional outcome measure in the economic analysis. This means that the difference in average total cost between both arms will be compared to the average difference in the PHQ-9 score between arms as an alternate assessment of value for money.

9.5. Analysis method

The statistical analyses for the economic evaluation will follow the principles detailed previously for the primary analyses. For the base case, generalised linear models (GLM) using a gamma family and log link will be used to estimate the difference in the total health sector costs between the intervention and control arms at 12 and 26 weeks. Separate GLMs will be used to estimate the difference in total societal costs between intervention and control arms at 12 and 26 weeks. Differences in QALYs between intervention and control arms at 12 and 26 weeks will be estimated with GLMs using gaussian family and identify link. All GLM models will adjust for baseline outcome measure.

In addition to reporting descriptive statistics and differences between both intervention and control arms for costs and outcomes, incremental cost-effectiveness ratios (ICERs) will be calculated. ICERs will be calculated as the mean difference in cost (from both health sector and societal) divided by the mean difference in the health outcome (QALYs, PHQ-9) between the study arms. Confidence intervals around the mean ICER will be estimated using a nested nonparametric bootstrap procedure within an imputation algorithm with 1000 iterations to reflect sampling uncertainty. The bootstrapped ICERs and the CIs will be graphically represented on cost-effectiveness planes and cost-effectiveness acceptability curves. A cost-effectiveness plane is a plot of the 1000 bootstrapped incremental costs and outcomes across four quadrants. A cost-effectiveness acceptability curve is a plot of the proportion of bootstrapped iterations that that fall below different willingness-to-pay values.

9.6. Sensitivity analyses

Sensitivity analyses will explore the effects of (1) using MBS/PBS data rather than self-reported data; and (2) the method to manage missing data.

10. Data management and workflow

Participant data are collected and stored in REDCap by blinded researchers during screening, by participants via electronic surveys, or by blinded researchers conducting general practice EMR audits.

Participants' EMR data will be collected from GP sites as part of a record audit to be performed by trial researchers. This data will be entered directly into REDCap alongside participant self-reported data. Administrative data regarding service use and prescription information from the Medical Benefits Scheme (MBS) and Pharmaceutical Benefits Scheme (PBS) will be sent securely to the University of Melbourne in compressed excel (.xls(x)) or comma separated value (.csv) format. Once all data has been collated, the REDCap database will be moved to analysis/cleanup status. All data

required for reporting and analysis of trial outcomes will be exported from REDCap as .csv files then imported into Stata Statistical Software v17.0 (or later) for review, cleaning, processing, and analysis.

The trial biostatistician will recode and remove the labels from the randomisation variable prior to any further work being undertaken with the data. An independent biostatistician not involved in trial activities and blinded to study arm status will check and clean the data, perform any manipulation of the raw data required for outcome derivation, and analyse the data according to the specifications contained in this SAP. Data checking and cleaning will involve verifying the validity of recorded values, renaming, and labelling of variables, deriving new variables, creating composite variables, and deleting any variables that are not required for analysis. Any errors will be corrected, and the nature of any potential/likely errors will be raised with trial researchers. The trial biostatistician will also work closely with the independent biostatistician to review the data coding and statistical analysis. The coding, derivation and analysis for the health economics outcomes will be conducted by the health economists.

10.1. Timing of final analysis and outcome assessment

Final statistical analysis is planned to occur after the signed SAP has been uploaded on the Australian and New Zealand Clinical Trials Registry, and all data has been collected, including the administrative data sources (PBS, MBS, and audit of the GP EMR) required for the analyses of the health economics outcome for first six-months of the trial period, and process evaluation.

End of data collection period for the primary and secondary trial outcomes will be at 26-weeks of follow-up post-randomisation. GP audit will commence soon after the last participant recruited at the first GP clinic reaches 26-weeks of follow-up and be completed after the last participant recruited reaches 26-weeks of follow-up. We expect that the MBS and PBS data for the first 6-months follow-up from baseline to be received late August 2024.

10.2. Longer term administrative data collection

PBS and MBS data between 7 to 12 months of follow-up that will be used for the analyses of the longer-term health economics outcome is expected to be requested and received in January 2025. The second tranche of PBS and MBS data (between 7 to 12 months of follow-up) will be requested and analysed PGx-informed prescribing is found to be effective or cost-effective at 26 weeks. This analysis will explore the potential duration of effect on resource use. Health economics outcomes will be analysed after a blinded review of the data has been conducted and reported separately to the primary and secondary outcomes.

11. Statistical software and technical details

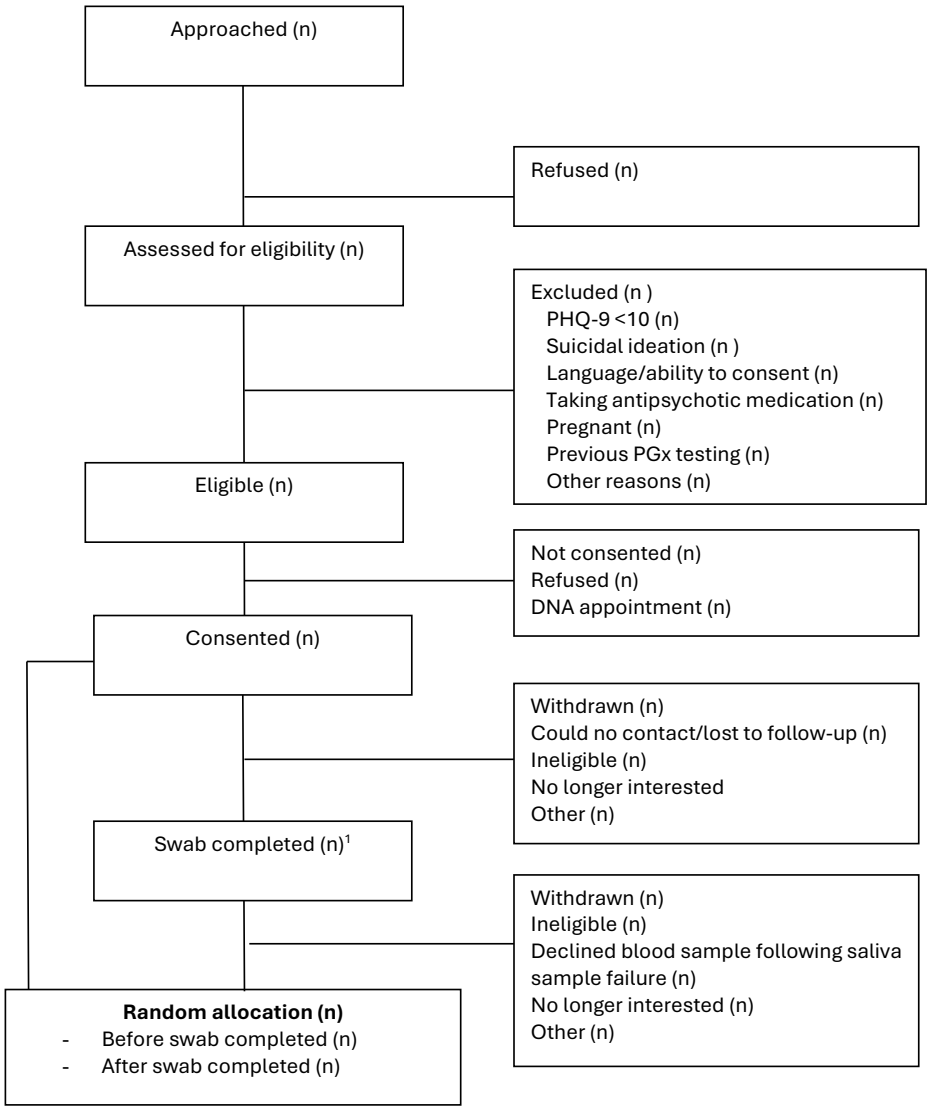
Data management and statistical analyses will be conducted using R[29] and Stata Statistical Software (v17)[30] or later. Appendix provides the proposed table shells for the presentation of demographics and baseline measures, and results of the statistical analysis of the primary and secondary outcomes. These results may also be presented graphically, where appropriate. Any post-hoc explanatory analyses not identified in the SAP will be clearly identified in the final statistical report. Any deviations from the planned analyses detailed in the SAP will be documented and reported in a revised version of this SAP.

12. References

1. Saya S, Chondros P, Abela A, et al. The PRESIDE (PhaRmacogEnomicS In DEpression) Trial: a double-blind randomised controlled trial of pharmacogenomic-informed prescribing of antidepressants on depression outcomes in patients with major depressive disorder in primary care. *Trials*. May 19 2023;24(1):342. doi:10.1186/s13063-023-07361-6
2. Kroenke K, Spitzer RL, Williams JB, Löwe B. The patient health questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. *General hospital psychiatry*. 2010;32(4):345-359.
3. Santomauro DF, Herrera AMM, Shadid J, et al. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *The Lancet*. 2021;398(10312):1700-1712.
4. Repišti S, Jovanović N, Kuzman MR, et al. How to measure the impact of the COVID-19 pandemic on quality of life: COV19-QoL–the development, reliability and validity of a new scale. *Global Psychiatry*. 2020;
5. Saveanu R, Etkin A, Duchemin A-M, et al. The international Study to Predict Optimized Treatment in Depression (iSPOT-D): outcomes from the acute phase of antidepressant treatment. *Journal of psychiatric research*. 2015;61:1-12.
6. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials–a practical guide with flowcharts. *BMC medical research methodology*. 2017;17:1-10.
7. Lee KJ, Carlin JB, Simpson JA, Moreno-Betancur M. Assumptions and analysis planning in studies with missing data in multiple variables: moving beyond the MCAR/MAR/MNAR classification. *International Journal of Epidemiology*. 2023;52(4):1268-1275.
8. Clark TP, Kahan BC, Phillips A, White I, Carpenter JR. Estimands: bringing clarity and focus to research questions in clinical trials. *BMJ open*. 2022;12(1):e052953.
9. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of biomedical informatics*. 2009;42(2):377-381.
10. Australian Bureau of Statistics. 2033.0.55.001 - Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australia, 2011. https://www.abs.gov.au/ausstats/abs@.nsf/lookup/2033.0.55.001main+features100052011
11. Australian Government Department of Health and Aged Care. Modified Monash Model. https://www.health.gov.au/topics/rural-health-workforce/classifications/mmm
12. Spitzer RL, Kroenke K, Williams JB, Group PHQPCS, Group PHQPCS. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *Jama*. 1999;282(18):1737-1744.
13. Lawrance R, Degtyarev E, Griffiths P, et al. What is an estimand & how does it relate to quantifying the effect of treatment on patient-reported quality of life outcomes in clinical trials? *Journal of Patient-Reported Outcomes*. 2020;4(1):1-8.
14. Fortney JC, Pyne JM, Edlund MJ, Mittal D. Relationship between antidepressant medication possession and treatment response. *General Hospital Psychiatry*. 2010;32(4):377-379.
15. Wisniewski SR, Rush AJ, Balasubramani G, Trivedi MH, Nierenberg AA, Investigators SD. Self-rated global measure of the frequency, intensity, and burden of side effects. *Journal of Psychiatric Practice®*. 2006;12(2):71-79.
16. Chan AHY, Horne R, Hankins M, Chisari C. The medication adherence report scale: a measurement tool for eliciting patients' reports of nonadherence. *British journal of clinical pharmacology*. 2020;86(7):1281-1288.
17. Ukoumunne OC, Forbes AB, Carlin JB, Gulliford MC. Comparison of the risk difference, risk ratio and odds ratio scales for quantifying the unadjusted intervention effect in cluster randomized trials. *Statistics in medicine*. 2008;27(25):5143-5155.
18. Pedroza C, Truong VT. Performance of models for estimating absolute risk difference in multicenter trials with binary outcome. *BMC medical research methodology*. 2016;16:1-12.

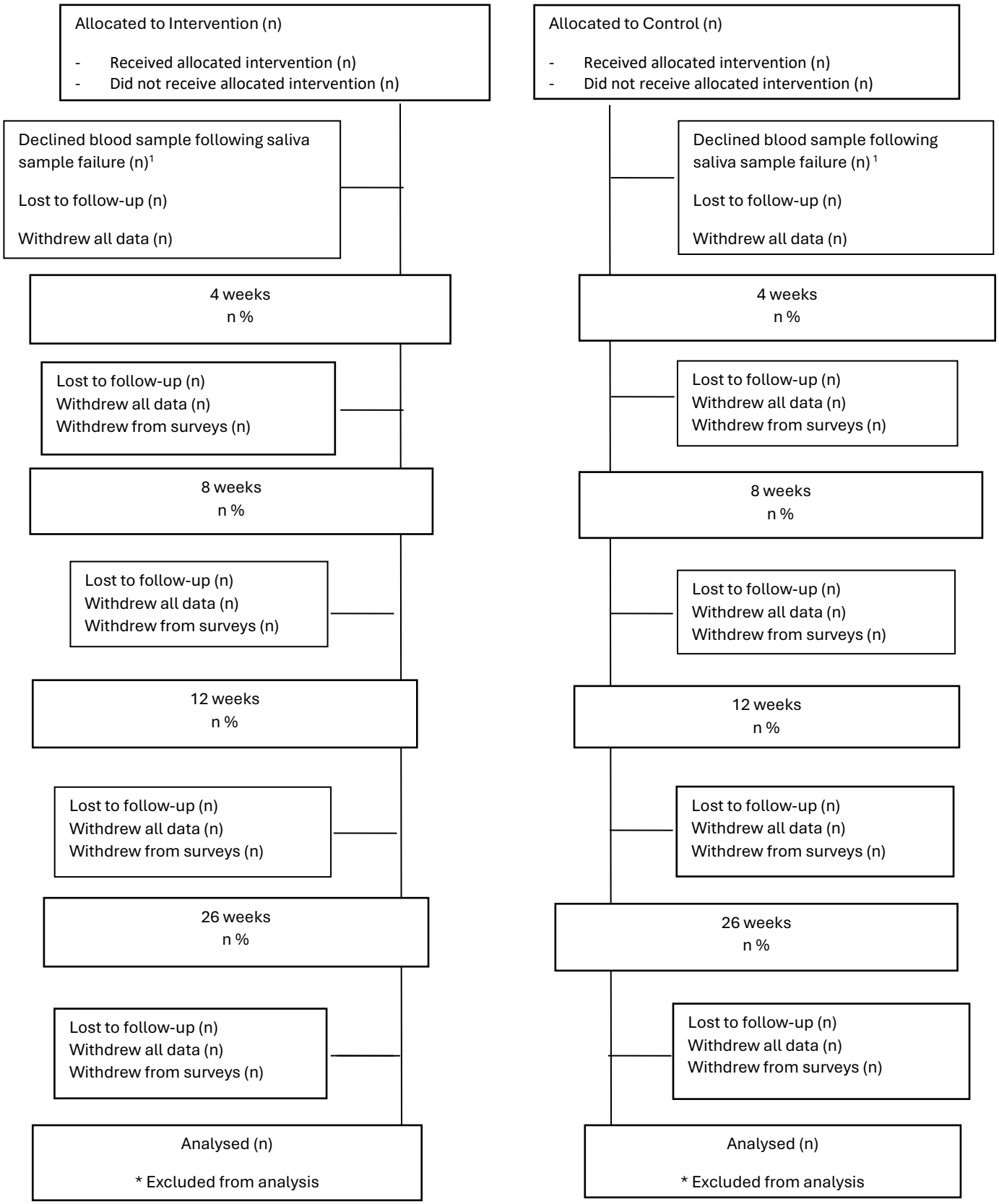
19. Dunn G, Maracy M, Tomenson B. Estimating treatment effects from randomized clinical trials with noncompliance and loss to follow-up: the role of instrumental variable methods. *Statistical methods in medical research*. 2005;14(4):369-395.
20. Hossain MB, Karim ME. Key considerations for choosing a statistical method to deal with incomplete treatment adherence in pragmatic trials. *Pharmaceutical Statistics*. 2023;22(1):205-231.
21. Vansteelandt S, Bowden J, Babanezhad M, Goetghebeur E. On instrumental variables estimation of causal odds ratios. 2011;
22. Braund TA, Tillman G, Palmer DM, Gordon E, Rush AJ, Harris AW. Antidepressant side effects and their impact on treatment outcome in people with major depressive disorder: an iSPOT-D report. *Translational psychiatry*. 2021;11(1):417.
23. Sonic Genetics. Pharmacogenomic (PGx) screen. https://www.sonicgenetics.com.au/our-tests/all-our-tests/pharmacogenomic-pgx-screen/
24. Australia AGS. Medicare Item Reports. http://medicarestatistics.humanservices.gov.au/statistics/mbs_item.jsp
25. Australian Government Department of Health and Aged Care. The Pharmaceuitcal Benefits Scheme Ex-manufacturer price. https://www.pbs.gov.au/info/industry/pricing/ex-manufacturer-price
26. Australian Bureau of Statistics. Average Weekly Earnings, Australia. https://www.abs.gov.au/statistics/labour/earnings-and-working-conditions/average-weekly-earnings-australia/may-2023
27. Jacobs P, Fassbender K. The measurement of indirect costs in the health economics evaluation literature: a review. *International journal of technology assessment in health care*. 1998;14(4):799-808.
28. Assessment of Quality of Life. Scoring - Psychometric (unweighted) or Utility (weighted)? https://aqol.com.au/index.php/scoring-algorithms
29. *R: A language and environment for statistical computing*. 2016.
30. *Stata Statistical Software: Release 17*. College Station, TX: StataCorp LLC.; 2021;

13. Appendices



[1] XX individuals randomised before completing the swab

Figure A1.1: PRESIDE consort diagram (Part 1: Enrolment)



[1] XX individuals in interventions and xx in control arm randomised before completing the swab

Figure A1.2: PRESIDE consort diagram (Part 2: Allocation, Follow-up and Analysis)

PRESIDE Proposed Tables for Primary, Secondary and Descriptive objectives

Table A1 - Participant demographics and baseline measures by study arm

	Intervention		Control		All participants	
	n	**(%)**	**n**	**(%)**	**n**	**(%)**
Depressive symptom severity (PHQ-9) – Mean (SD)						
Health-related quality of life (AQoL-4D utility score) – Mean (SD)						
Age (years) – Mean (SD)						
Gender						
Female						
Male						
Other						
Born in Australia						
English spoken at home						
Ethnicity[1]						
European[2]						
Aboriginal and/or Torres Strait Islander						
Central/South Asian[3]						
Near Eastern[4]						
East Asian[5]						
Oceanian[6]						
Latin American						
Sub-Saharan African						
Other						
Index of Relative Socio-economic Disadvantage (IRSD quintiles) for participant residence						
Most disadvantaged 1						
2						
3						
4						
Least disadvantaged 5						
Highest level of education completed						
Less than Year 10						
Year 10						
Year 11						
Year 12 or equivalent						
Certificate III/IV						
Advanced Diploma/Diploma						
Bachelor's Degree						
Graduate Diploma/Graduate Certificate						
Postgraduate Degree						
Usual weekly employment						
Working for an employer for wages/salary						
Working in your own business for profit/pay						
Working without pay in a family business or farm						
Unemployed, looking for work and available to start						
Unemployed, unable to work						

	Intervention	Control	All participants
	n (%)	n (%)	n (%)
None of the above			
Lives alone			
Living arrangements if do not live alone[1]			
Husband or wife			
Defacto partner			
My child/ren			
My partner's child/ren			
My parent/s			
Unrelated flatmate or co-tenant			
Other			
Current AD status			
Prescribed			
Not Prescribed			
Not sure			
Previous psychosis			
Experienced psychosis previously			
Never experienced psychosis			
Not sure			

SD – Standard deviation; Data presented are counts (n) and column percentages (%) unless labelled otherwise.

[1]Participants were able to tick all boxes that applied, so percentages may not sum to 100%, [2]European Australian, U.K., Greece, France, Germany, Spain, Italy, [3] Pakistan, Sri Lanka, Bangladesh, India, [4] Northern Africa, Middle East, Turkey, [5] Japan, Korea, China, [6]Hawaiian, Papua New Guinea

Note: In the publication, additional variables collected at baseline may also be presented in the table. Categorical variables may be collapsed in the tables and statistical analysis.

Table A2.1 Change in depressive symptom score (PHQ-9) at each time point by study arm

	Intervention		Control				
Time-point	**n**	**Mean (SD)**	**n**	**Mean (SD)**	**Difference[1]**	**95% CI[1]**	**p-value**
12 weeks (primary outcome)							
Primary Analysis[2]							
Sensitivity analysis[3,]							
Sensitivity analysis[4]							
Sensitivity analysis[5]							
Sensitivity analysis[6]							
Adherence adjusted analysis[7]							
4 weeks (secondary outcome)							
Primary Analysis[2]							
Sensitivity analysis[3]							
Sensitivity analysis[4]							
Sensitivity analysis[5]							
Sensitivity analysis[6]							
Adherence adjusted analysis[7]							
8 weeks (secondary outcome)							
Primary Analysis[2]							
Sensitivity analysis[3]							
Sensitivity analysis[4]							
Sensitivity analysis[5]							
Sensitivity analysis[6]							
Adherence adjusted analysis[7]							
26 weeks (secondary outcome)							
Primary Analysis[2]							
Sensitivity analysis[3]							
Sensitivity analysis[4]							
Sensitivity analysis[5]							
Sensitivity analysis[6]							
Adherence adjusted analysis[7]							

n – count; SD – Standard deviation; CI – Confidence interval

[1] Difference in mean between the intervention and control arms with respective 95% CI and p-value estimated using constrained linear mixed effects model with study arm, general practice, baseline antidepressant use, and time (baseline, four, eight, 12, and 26 weeks) included as fixed effects and random intercepts for individuals, with two-way interaction between study arm and time, except for baseline where study arm means were constrained to be equal.

[2] The primary outcome is the mean change at 12 weeks from baseline. Secondary outcomes are the mean change at 4, 8 and 26 weeks, from baseline

[3] Sensitivity analysis: Adjustment for additional covariates

[4] Sensitivity analysis: Survey outcome responses are outside the assessment window set to missing

[5] Sensitivity analysis: Analysis using hypothetical approach for intercurrent event related to the initiation of antipsychotic medications

[6] Sensitivity analysis: Analysis for missing data may also be included based on blinded review of the missing data patterns and mechanisms.

[7] Adherence adjusted analysis: Analysis including binary indicator variable for adherence (see Section 8.4.3)

Table A2.2 Sub-group analysis for change in depressive symptom score (PHQ-9) at each time point by study arm for (i) AD Medications use at baseline, and (ii) Depressive symptom severity at baseline

	Intervention		Control				
Time point	**n**	**Mean (SD)**	**n**	**Mean (SD)**	**Difference[1]**	**95% CI[1]**	**p-value**
Baseline AD medication use							
4 weeks							
Never taken							
Prior AD use but not currently taking							
Currently taking							
8 weeks							
Never taken							
Prior AD use but not currently taking							
Currently taking							
12 weeks							
Never taken							
Prior AD use but not currently taking							
Currently taking							
26 weeks							
Never taken							
Prior AD use but not currently taking							
Currently taking							
Baseline depressive symptom severity							
4 weeks							
Moderate (PHQ-9: 10 to 14)							
Moderately severe (PHQ-9: 15 to 19)							
Severe (PHQ-9: 20 to 27)							
8 weeks							
Moderate (PHQ-9: 10 to 14)							
Moderately severe (PHQ-9: 15 to 19)							
Severe (PHQ-9: 20 to 27)							
12 weeks							
Moderate (PHQ-9: 10 to 14)							
Moderately severe (PHQ-9: 15 to 19)							
Severe (PHQ-9: 20 to 27)							
26 weeks							
Moderate (PHQ-9: 10 to 14)							
Moderately severe (PHQ-9: 15 to 19)							
Severe (PHQ-9: 20 to 27)							

n – count; SD – Standard deviation; CI – Confidence interval

[1] Difference in mean between the intervention and the control arms with respective 95% CI and p-value estimated using constrained linear mixed effects model with study arm, general practice, baseline antidepressant use, and time (baseline, 4, 8, 12, and 26 weeks) included as fixed effects and random intercepts for individuals, with three-way interaction between the sub-group variable, study arm and time, except for baseline where study arm means were constrained to be equal.
[2] The primary outcome is the mean change at 12 weeks from baseline. Secondary outcomes are the mean change at 4, 8 and 26 weeks from baseline
[3] Effect modification at each time point by (1) participant baseline AD use (p-values for interaction effect at: 4 weeks X.XXX; 8 weeks X.XXX; 12 weeks X.XXX; 26 weeks X.XXX), and (2) baseline depressive symptom severity (p-values for interaction effect at: 4 weeks X.XXX; 8 weeks X.XXX; 12 weeks X.XXX; 26 weeks X.XXX)

Table A3 – Depressive symptom status at 12-week follow-up by study arms

	Intervention n (%)	Control n (%)	Diff (95% CI)[1]	OR (95% CI)[2]	p-value[3]
Depressive symptom remissions					
Primary Analysis					
Sensitivity analysis[3]					
Sensitivity analysis[4]					
Sensitivity analysis[5]					
Sensitivity analysis[6]					
Adherence adjusted analysis[7]					
Depressive symptom response					
Primary Analysis					
Sensitivity analysis[3]					
Sensitivity analysis[4]					
Sensitivity analysis[5]					
Sensitivity analysis[6]					
Adherence adjusted analysis[7]					

n – count; OR – Odds Ratio; CI – Confidence Interval

[1]Difference in the percentage and respective 95% CI between the intervention and control arms estimated using generalised linear model with the identity link function and binomial family adjusted for general practice, baseline antidepressant use, and baseline PHQ-9 score.

[2]Odds ratio of the intervention arm compared to the control arm and respective 95% CI estimated using logistic regression adjusted for general practice, baseline antidepressant use, and baseline PHQ-9 score.

[3]P-values presented are associated with the estimated odds ratio.

[3] Sensitivity analysis: Adjustment for additional covariates

[4] Sensitivity analysis: Survey outcome responses are outside the assessment window set to missing

[5] Sensitivity analysis: Analysis using hypothetical approach for intercurrent event related to the initiation of antipsychotic medications

[6] Sensitivity analysis: Analysis for missing data may also be included based on blinded review of the missing data patterns and mechanisms.

[7] Adherence adjusted analysis: Analysis including binary indicator variable for adherence (see Section 8.4.3)

Table A4 – Antidepressant side-effects and medication adherence by study arm

Time point	Intervention		Control	
	n	Mean (SD)	n	Mean (SD)
Antidepressant side-effects (FIBSER)[1]				
4 weeks				
8 weeks				
12 weeks				
26 weeks				
Adherence to prescribed antidepressants (MARS-5)[1]				
4 weeks				
8 weeks				
12 weeks				
26 weeks				

Table A5 - Medication procession ratio and antidepressant medication changes by study arm

	Intervention	Control	Estimated effect size (95% CI)		
Medication possession ratio (MPR)[1]	N MPR	N MPR	Difference in MPR	Ratio of MPR	p-value
Sensitivity Analysis[3]			Difference in MPR	Ratio of MPR	p-value
Sensitivity Analysis[4]			Difference in MPR	Ratio of MPR	p-value
Sensitivity Analysis[5]			Difference in MPR	Ratio of MPR	p-value
Rate in antidepressant medication changes[2]	N rate	N rate	Difference in rate	Rate Ratio	p-value
Sensitivity Analysis[3]			Difference in rate	Rate Ratio	p-value
Sensitivity Analysis[4]			Difference in rate	Rate Ratio	p-value
Sensitivity Analysis[5]			Difference in rate	Rate Ratio	p-value

N-Number of participants; MPR-total number of days possessed AD medication/total number of days observed; rate - total number of AD changes/total number of days observed.

[1] Difference in MPR and ratio of MPR between the intervention arm compared to the control arm with respective 95% CI and p-value estimated using Poisson regression adjusted for general practice, baseline antidepressant use and baseline PHQ-9 score.

[2] Difference in the rates and rate ratio of the intervention arm compared to the control arm with respective 95% CI estimated and p-value using Poisson regression adjusted for general practice, baseline antidepressant use and baseline PHQ-9 score.

[3] Sensitivity analysis: Adjustment for additional covariates

[4] Sensitivity analysis: Analysis using hypothetical approach for intercurrent event related to the initiation of antipsychotic medications

[5] Sensitivity analysis: Analysis for missing data may also be included based on blinded review of the missing data patterns and mechanisms.

NOTE: Similar tables will be created to report the analysis for participants who have an actionable drug-gene combination (See additional estimands, section 8.4.3) and health economic analysis (Section 9).