Statistical analysis plan for Link-me: A randomised controlled trial of a systematic approach to stepped mental health care in primary care

Date and version number

31.05.2019 Version 1

Trial registration

Link-me was prospectively registered on the Australian and New Zealand Clinical Trials Registry (ACTRN 12617001333303) in September 2017. Participant recruitment commenced in November 2017 and was completed in October 2018. Follow-up and process data collection will be completed 13 months after the last date of patient's recruitment, allowing for an additional month to follow-up late responders to 12 month outcome measures.

Funding acknowledgement

This trial was conducted as part of the national evaluation of the Primary Health Network Mental Health Reform Lead Site Project funded by the Australian Government Department of Health. As part of the funding requirements a report of the trial outcomes will be submitted to the Department of Health in July 2019. Link-me builds on the *diamond* and Target-D studies both funded by the National Health and Medical Research Council (NHMRC) (IDs: 299869, 454463, 566511, 1002908, and 1059863). We acknowledge the many dedicated GPs, general practice staff, and patients for their commitment to these studies and for making the current trial possible. Thanks also to the Primary Health Network Lead Sites and Department of Health for their contribution to the development and implementation of the Link-me trial.

Protocol reference

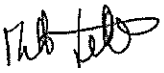
Fletcher S, Chondros P, Palmer VJ, Chatterton ML, Spittal MJ, Mihalopoulos C, Wood A, Harris M, Burgess P, Bassilios B, Pirkis J, Gunn J. Link-me: Protocol for a randomised controlled trial of a systematic approach to stepped mental health care in primary care. Contemporary Clinical Trials 2019; 78: 63-75. DOI: 10.1016/j.cct.2018.12.014

SAP revision history

Protocol version	Updated SAP version no.	Section number changed	Description and reason for change	Date changed
Version 1				

Roles and responsibilities

Signatory names and contribution

Name	Role	Affiliation	Signature
Matthew Spittal*	Senior statistician	Melbourne School of Population and Global Health, University of Melbourne, VIC, Australia **e-mail**: m.spittal@unimelb.edu.au	
Patty Chondros	Senior statistician	Department of General Practice, University of Melbourne, Melbourne, VIC, Australia	
Cathrine Mihalopoulos	Lead health economist	School of Health and Social Development, Deakin University, Burwood, VIC, Australia	
Susan Fletcher	Project manager	Department of General Practice, University of Melbourne, Melbourne, VIC, Australia	

*Contact person

Non-signatory names and contribution

Name	Role	Affiliation
Mary Lou Chatterton	Health economist	School of Health and Social Development, Deakin University, Burwood, VIC, Australia
Konstancja Densley	Data manager and analyst	Department of General Practice, University of Melbourne, Melbourne, VIC, Australia
Meredith Harris	Health services researcher	School of Public Health, The University of Queensland, QLD, Australia
Philip Burgess	Health services researcher	School of Public Health, The University of Queensland, QLD, Australia
Bridget Bassilios	Investigator	Melbourne School of Population and Global Health, University of Melbourne
Victoria Palmer	Investigator	Department of General Practice, University of Melbourne, Melbourne, VIC, Australia
Jane Pirkis	Chief investigator	Melbourne School of Population and Global Health, University of Melbourne, Melbourne, VIC, Australia
Jane Gunn	Chief investigator/ Clinical lead	Department of General Practice, University of Melbourne, Melbourne, VIC, Australia

Abbreviations

AR-DRGs	Australian-refined diagnostic related groups
CACE	Complier average causal effect
CI	Confidence interval
CSV	Comma separated values
DSM	Diagnostic and Statistical Manual
DST	Decision Support Tool
EQ-5D-5L	Euroqol 5-dimension quality of life questionnaire (5-level version)
GAD-2 / GAD-7	Generalised Anxiety Disorder scale (2-item / 7-item version)
GLM	Generalised linear models
GP	General practitioner
ICER	Incremental cost-effectiveness ratios
IHPA	Independent Hospital Pricing Authority
IQR	Interquartile range
ITT	Intention to treat
K10 / K10+	Kessler Psychological Distress Scale (10-item / 14-item version)
MAR	Missing at random
MBS	Medicare Benefit Schedule
NWAUs	National Weighted Activity Units
PBS	Pharmaceutical Benefit Scheme
PHN	Primary Health Network
PHQ-2 / PHQ-9	Patient Health Questionnaire (2-item / 9-item version)
PMHC MDS	Primary Mental Health Care Minimum Data Set
QALY	Quality adjusted life year
RCT	Randomised controlled trial
RUQ	Resource Use Questionnaire
SD	Standard deviation
VAS	Visual analogue scale

Contents

Background

Link-me is a stratified individually randomised controlled trial (RCT) with general practice patients that aims to determine whether using the Decision Support Tool (DST) to systematically identify general practice patients' predicted symptom severity for depression and/or anxiety and provide tailored treatment recommendations is clinically and cost effective compared to usual care. The study rationale, and details of the study design, including setting, eligibility criteria, sample size calculations and statistical analysis are detailed in the published study trial protocol [1]. **This document provides a detailed statistical analysis plan, including the economic evaluation, to complement the study protocol and to expand on the secondary and sensitivity analyses.**

A separate protocol will be developed for the process evaluation to be conducted in parallel with the randomised controlled trial. The process evaluation will provide the context to help understand the outcomes that were achieved, identify challenges in implementation and provide important guidance for future translation of trial findings using the framework set out by the Medical Research Council [2].

Primary hypothesis

The null hypothesis is that there is no difference in mean psychological distress scores between the intervention and comparison arms at six months. The alternative hypothesis is that there is a difference between the two trial arms.

The primary objective of the trial is to determine whether there is a difference in mean psychological distress scores between the intervention and comparison arms at six months.

Secondary objectives of the trial are to determine

a) whether there is a difference in mean psychological distress at 12 months between the two trial arms
b) whether there is a difference in quality of life, days out of role, depressive symptoms, anxiety symptoms, self-efficacy, health service use and cost between trial arms, at six and 12 months, respectively
c) whether the primary and secondary outcomes differ between the two trial arms within each of the mild/minimal and severe symptom groups; and
c) the cost-effectiveness of the new model of care compared to usual care.

Trial methods

In brief, Link-me is set in general practices located within the three participating Primary Health Network (PHN) catchment areas, across three states in Australia (Victoria, New South Wales and Queensland). At least 18 general practices will participate in the trial (minimum of six in each PHN). Patients are approached in the general practice waiting room and assessed for eligibility using a self-report survey on a hand-held tablet device. Patients are eligible if they are aged between 18 and 75, able to complete the survey in English, provide a phone number and email address, and hold a Medicare card, and meet one or more of the following criteria:

- a score of two or more on the 2-item version of the Patient Health Questionnaire [PHQ-2:3]
- a score of two or more on the 2-item version of the Generalised Anxiety Disorder scale [GAD-2:4];
- current use of medication for mental health problems.

Those who are eligible and consenting are then invited to complete the Link-me DST. The Link-me DST comprises 23 items assessing current depressive symptoms, current anxiety symptoms, lifetime history of depression, gender, living situation, ability to manage on available income, self-rated general health, and presence of chronic illness that affects the ability to carry out daily activities. Responses are used in two prognostic models embedded within the DST to predict symptom scores for anxiety and depression at three months. Based on their predicted score, participants are classified into one of three symptom severity groups (minimal/mild, moderate, and severe).

Participants who are assessed as having "minimal/mild" or "severe" symptoms will be randomised to the intervention or comparison arms and will receive information relevant to their symptom severity group and trial arm allocation. Patients allocated to the intervention arm receive feedback on DST responses, select treatment priorities, assess motivation to change, and receive a severity-matched treatment recommendation (information about low intensity services for those with mild symptoms, or assistance from a specially trained health professional ['care navigation']for those with severe symptoms). All patients allocated to the comparison arm receive usual GP care plus attention control.

Participants with moderate symptoms will not be randomised as they are assumed to be appropriately served by the existing mental health service options available via their GP. Although they are not part of the Link-me trial, they will be followed up and asked to complete outcome assessments at six and 12 months. Data on the participants classified as moderate together with the trial data will be used to update and validate the *diamond* clinical prediction tool [5]. Details of the secondary use of the trial data will be described in separate statistical analysis plans.

Participant recruitment commenced on 21 November 2017 and was completed on 31 October 2018. Six and 12 month follow-up assessments will be completed for all participants by 12 June 2019 and 11 December 2019 respectively.

Outcomes

Outcomes are assessed at trial enrolment and at six and 12 months after DST completion.

The **primary outcome** is psychological distress as measured using the 10 standard items of the Kessler Psychological Distress Scale [K10:6]. Respondents are asked to indicate how often in the past four weeks they have experienced certain symptoms (e.g., nervousness, hopelessness, fatigue, agitation, and depressed mood), using a five-point Likert scale (where 1='not at all' and 5='all the time'). The total K10 score is the sum of the 10 items, ranging from 10 to 50, where higher K10 scores indicate greater higher psychological distress [7]. If two or fewer items on the K10 are missing responses, the missing values will be substituted with the mean response of the completed items, otherwise the total score will be coded as missing.

K10 psychological distress at 12 months is a **secondary outcome.** Other secondary outcomes are depressive symptom severity, anxiety symptom severity, quality of life (utility scores and overall health), and days out of role, all assessed at both six and 12 months.

Depressive symptom severity will be measured using the 9-item version of the Patient Health Questionnaire [PHQ-9:8] which assesses the presence of the nine Diagnostic and Statistical Manual (DSM) symptoms of depression over the last two weeks using a four-point Likert scale (0=Not at all, 1=Several days, 2=More than half the days, 3=Nearly every day). Total scores are calculated by summing the nine items, and range between zero and 27. Low scores (0-9) indicate minimal/mild

symptoms for depression and higher scores (>20) indicate severe depressive symptoms. If two or fewer items on the PHQ-9 are missing responses, the missing values will be substituted with the mean response of the completed items, otherwise the total score will be coded as missing [9].

Anxiety symptom severity will be measured using the seven-item Generalized Anxiety Disorder scale [GAD-7: 10]. The GAD-7 assesses the presence of generalised anxiety symptoms over the past two weeks using a four-point Likert scale (0=Not at all, 1=Several days, 2=More than half the days, 3=Nearly every day). The seven items are added to create a total score, which ranges between zero and 21, where the higher scores indicate more severe anxiety symptoms. If one or two items on the GAD-7 are missing a response, the missing values of these items will be substituted with the mean response of the completed items, otherwise the total score will be coded as missing.

Days out of role will be measured with two of the additional items from the K10+ [11], namely "In the last four weeks, how many days were you totally unable to work, study, or manage your day to day activities because of these feelings?**"** (days totally out of role) and **"**Aside from those days, in the last four weeks, how many days were you able to work, study, or manage your day to day activities, but had to cut down on what you did because of these feelings?" (days partially out of role) . The K10+ is a four-item extension of the K10 that respondents are asked to complete when they score greater than 10 on the standard 10-item scale (K10). In addition to the two items assessing days totally and partially out of role as above, one item asks respondents to indicate the number of health professional consultations sought as a result of their psychological distress in the past four weeks, and the extent to which physical health problems were the main cause of distress. For participants with a K10 score less than or equal to 10, the values for days totally or partially out of role will be coded as zero as participants with no symptoms of psychological distress can be reasonably expected to have had no days off or limited ability to work their usual role for this reason.

Quality of life will be measured using the utility index and the visual analogue scale (VAS) from the EQ-5D-5L, a self-report scale assessing health states across five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) [12]. Respondents rate the extent to which they have problems in each dimension on a five-point scale, and indicate their overall health on a VAS scale from zero to 100. The utility index will be calculated using the EuroQol group's guidelines [13] and an Australian valuation data set from [14].

A self-report resource use questionnaire (RUQ) will be used to collect health service use as part of the economic evaluation. To minimise participant burden, the RUQ is only administered at six and 12 months. Health service use at baseline will be assessed using one of the additional four items on the K10+ (see above) and routinely collected data about health service use for those who consent, assuming that most highly used services are captured in these datasets. Note that Medicare Benefits Schedule (MBS) and Pharmaceutical Benefits Scheme (PBS) data will only be included in the analysis of 12-month data. This is due to the complexity and time lag involved in the process required for data extraction from the Australian Government Department of Human Services. Details of the outcome measures are provided in the Link-me study protocol.

Screening and baseline data collection

General practice and general practitioner (GP) characteristics are collected at the time of practice recruitment, prior the commencement of patient recruitment in each practice. Data collected include (but are not limited to):

General practice characteristics

- Practice location (relative socioeconomic advantage/disadvantage of area [15])
- Practice type (private, corporatised, other)
- Practice size (staff head count and FTE including GPs, nurses, mental health professionals, allied health, other)
- Billing model (private or bulk billing)
- Co-location and record sharing with psychologist, counsellor, nurse, other)

General practitioner characteristics

- Age (years)
- Gender (male, female)
- Country of Graduation (Australia, overseas)
- Years in general practice (in Australia and overseas)
- Proportion of consultations conducted in English or languages other than English (with/without interpreter)
- Usual approach to mental health care (e.g. number of contacts with case manager, use of strategies e.g., conduct standardised assessment, recommend self-help, prescribe medication, refer).

Participant characteristics

Table 1 below outlines the patient information collected during the enrolment process, including the item wording and response options.

Table 1. Demographic characteristics collected at screening for individuals in their GP waiting room

Item description	Questions in screening survey	Responses
Age in years[1]	What is your age (in years)?	Free text number field (valid responses 0-99, inclusive)
Gender[2]	What is your gender?	Male Female Other
Indigenous status	Are you of Aboriginal or Torres Strait Islander origin?	Aboriginal Torres Strait Islander Aboriginal and Torres Strait Islander None of the above
Language mainly spoken at home	Which language do you mainly speak at home?	English Other
Highest level of education completed	What is the highest level of education you have completed?	Below Year 10 Year 10 / equivalent Year 11 / equivalent Year 12 / equivalent Certificate III/IV Advanced diploma / Diploma Bachelor's degree or higher Graduate diploma/Certificate Postgraduate degree

Item description	Questions in screening survey	Responses
Current employment	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 None of the above
Main activity for those not working or looking for work	In a usual week, which of the following best describes your main activity?	Retired or voluntarily inactive Home duties Caring for children Studying Unable to work due to own illness, injury, or disability Caring for an ill or disabled person Working in an unpaid voluntary job Other
Holds a health care card	Do you currently hold an Australian Government Health Care Card or Pensioner Concession Card?	Yes No
Depressive symptom severity (PHQ-2)[2,3]	Over the ***last 2 weeks***, how often have you been bothered by little interest or pleasure in doing things? Over the ***last 2 weeks***, how often have you been bothered by feeling down, depressed or hopeless?	For each item the responses are: 0=Not at all 1=Several days 2=More than half the days 3=Nearly every day Total score is the sum of two items
Anxiety symptom severity (GAD-2)[2,3]	Over the ***last 2 weeks***, how often have you been bothered by feeling nervous, anxious, or on edge? Over the ***last 2 weeks***, how often have you been bothered by not being able to stop or control worrying?	For each item the responses are: 0=Not at all 1=Several days 2=More than half the days 3=Nearly every day Total score is the sum of two items
Currently taking medication for mental health[2,3]	Do you take any medication for your mental health?	Yes No

1 Item assesses eligibility for Link-me: Patients ineligible and exit survey if age<18 or >75 years
2 Item included in the DST
3 Item assesses eligibility for Link-me: Patients ineligible and exit survey if PHQ-2<2 and GAD-2<2 and not currently taking medication for mental health

Eligible and consenting participants then completed the baseline survey, which in addition to the outcome measures described under 'Outcomes' on pages 6-7, included items assessing reason for

GP visit, general health, living situation, financial stability, and depression history, as detailed in Table 2 below.

Table 2. Characteristics collected at baseline, for eligible individuals that consented to participate in the trial

Item description	Questions in baseline survey	Responses
Reason for visit to GP	Is your visit to the doctor today mainly related to your	Physical health Mental health and wellbeing Both physical and mental health None of these
Long-term illness which limits daily activities[1]	Do you have any long-term illness, health problem, which limits your daily activities or the work you can do (including problems that are due to old age)?	Yes No
Self-rated health[1]	In general, would you say your health is…	Excellent Very good Good Fair Poor
Live alone[1]	Do you live alone?	Yes No
Managing on your available income[1]	How do you manage on your available income?	Easily Not too bad Difficult some of the time Difficult all of the time Impossible
History of depression[1]	Have you ever been bothered by feeling down, depressed or hopeless for longer than 2 weeks? Have you ever been bothered by little interest or pleasure in doing things for longer than 2 weeks?	For each item the responses are: Yes No Combined responses of the two items to create a new binary variable: 1 if responded yes to both items and 0 (no) otherwise

[1] Item included in the DST

Data management and workflow

Data collection and management processes are described in the trial protocol. This section describes how the data collected via a purpose-built secure online data collection system (developed and maintained by Strategic Data) will be prepared for statistical analyses. At the conclusion of each data collection phase (31 October 2018 (baseline), 12 June 2019 (6 months), and 11 December 2019 (12 months)), the project manager will contact Strategic Data and request the close of the relevant

survey and for all data for that phase to be extracted. Strategic Data will then de-identify the trial data within the online system using unique record identifiers for each participant and extract this de-identified data from the online data collection system in the form of comma separated text (CSV) data files. A unique link to these files will be emailed to the project manager within two working days of the data extraction request, and a password to access them sent separately to the project manager via SMS. This password will remain valid for 24 hours. The project manager will then download the CSV data files and substitute trial arm allocation for a dummy coded variable (A/B) and store the key to this code in a locked filing cabinet. The project manager will then save the dummy-coded files to the central password-protected University system where they will be stored securely and backed up regularly.

The data manager, blinded to trial arm allocation, will then import the CSV files into Stata [16] for data processing and statistical analysis. Data will be checked to identify and where possible resolve errors prior to analyses being conducted. Steps will include labelling the variables and values, creating composite variables and creating the total scores according to the instrument's guidelines. Datasets from each data collection point (screening/baseline, 6 and 12 months) will be merged using participant's generated unique identifier (ID). De-identified data will be stored on the University server for future use in accordance with the National Statement on Ethical Conduct in Human Research [17][1].

Harms

The trial protocol (Section 2.11.1) describes how the harms and safety concerns will be monitored and managed. No interim analyses or auditing are planned to determine harm from the intervention directly because the services recommended in the trial (both low intensity and those identified during care navigation) are also available to individuals outside the trial, and all participants are linked in with health services. Any adverse events reported (such as high levels of suicidal ideation) will be summarised using counts and percentages by trial arm and symptom severity group.

Statistical methods

Statistical analyses will be conducted at the end of the six and 12-month data collection periods separately, with the six months data being used for the primary analysis. In order to meet funder requirements, 6-month data analysis will be conducted blinded in June and July 2019, and a draft and final report of the findings will be submitted to the Australian Government Department of Health by 31 July and 30 September 2019, respectively. Twelve-month data analysis will commence in early 2020 with the final report to be submitted to the Department of Health by 30 September 2020. No interim analyses are planned. All analysis will be conducted using Stata [16], and, if required, the R statistical package [18].

The results may also be submitted to peer-reviewed journals for publication and presented at conferences, contingent on approval from the Department of Health.

[1] Note that this data retention plan differs from that initially approved by the Human Research Ethics Committee and described in the published protocol due to the release of new guidance by the National Health and Medical Research Council in 2018.

Descriptive analysis

A flow chart will be created to show the participant flow through the study (template in Figure 1). The flow chart will show the recruitment rate (including the number of participants screened, eligible, consented), the number randomised to the trial arms, attrition rates, and the number of participants that contributed outcome data at each assessment time point by trial arm.

Data collected in the screening survey will be used to describe the number of patients that were ineligible, noting that patients are exited from the survey as soon as they are identified as ineligible; either in response to the first question (aged <18 years or >75 years old) or at the end of the survey (both the PHQ-2 and GAD-2 scores <2 and patient does not report current medication for mental health). Therefore, where patients indicate an age outside the eligible range, they are not asked to complete the rest of the survey so can not be included in subsequent descriptive analysis.

Attrition rates and number of participants with outcome assessments at each time point (by trial arm) will also be reported for each symptom severity group. When such information is available, the reasons participants withdrew or lost to follow up will be reported by trial arm and symptom severity group.

Descriptive statistics will be used to summarise socio-demographic and clinical characteristics of participants collected at screening and baseline by trial arm, for the total sample and by each symptom severity group as shown in Table 1. For continuous data with a skewed distribution, medians and quartiles will be used instead.

Primary analysis

Primary analysis will be based on an intention-to-treat (ITT) approach [19], where all individuals randomised will be included in the analysis by their allocated trial arm status regardless of whether they received all, part or none of the intended treatments.

For the primary analysis, linear regression will be used to estimate the difference in mean chance from baseline in the mean K10 psychological distress scores between the intervention and comparison arms at six months with adjustment for symptom severity group (minimal/mild vs. severe) and baseline K10 scores. No other baseline variables will be considered for adjustment in the analysis [20].

Although randomisation was stratified by general practice site, it will not be included as a fixed effect factor in the regression model for the primary analysis. The estimated treatment effects will be unbiased without the adjustment for general practice as the participants from each general practice will be balanced between the trial arms. Provided the general practice site was correlated with the outcome, there could be potential gains in precision for the estimated effect size if we included this in the analysis. However, given that patients will be recruited from a minimum of 18 general practices, the gains in precision of adjusting for the stratification factors (general practice) will potentially be outweighed by loss of precision in the estimates due to the large number of degree of freedoms lost for estimating many fixed effects for general practice [21]. However, as part of a sensitivity analysis, we will include general practice as a random effect in the complete case analysis (see below).

Multiple imputation will be used to handle the incomplete data. We will impute 50 datasets for six-month outcome data using chained equations to generate imputed data [22]. Predictor variables will be the primary and secondary outcomes measured at each assessment time (that is, baseline and six

months), plus selected baseline variables (trial arm status, symptom severity group, general practice, age and sex. All parameters of interest (e.g., means, mean differences, rate ratios) and their standard errors will be combined using standard methods for this type of data [i.e., Rubin's rules: 23].

As described above, the trial arm status will be coded with the letters A and B, and the key kept by the project manager to ensure that the study investigators, together with the lead statistician and health economist who will conduct and interpret the statistical and economic analysis, will remain blinded to the trial arm status of participants until the ITT analysis is conducted and interpreted.

Sensitivity analyses: Complete-case analysis and adjustment for covariates

Sensitivity analyses will be conducted using complete-cases only in analyses with and without adjustment for general practice effects. Complete cases will be defined as records where the outcome data was observed at baseline and at six months.

For the analysis with no adjustment for general practice, linear regression will be used to estimate the difference in mean change from baseline in the mean K10 scores between the intervention and comparison arms at six months with adjustment for symptom severity group (minimal/mild vs. severe) and baseline K10 scores.

For the analysis with adjustment for practice effects, we will use linear mixed effects regression with general practice treated as random intercept term, and symptom severity group (minimal/mild vs. severe) and baseline K10 scores as fixed effects.

Sub-group analyses

Sub-group analyses conducted in the mild/minimal and severe symptom groups will estimate, for each sub-group, change from baseline in mean K10 scores at six months between the two trial arms, adjusting for baseline K10 scores and using multiple imputation. We will also undertake sensitivity analyses for the sub-group analyses using complete-case data and also adjusting for practice effects using a random intercept model.

Secondary outcomes

We will repeat all analyses described above for the primary outcome for the 12-month K10 scores (that is, the primary analysis and the two sensitivity analyses). The multiple imputations will be repeated for the 12-month data using the same approach described for the six-month data, except the imputation model will include all the variables included for the six-month imputed data plus the outcome measures measured at 12 months. For the complete case analysis of the 12 month outcomes, complete cases will be records with outcome data observed at baseline and 12 months.

We will also undertake analyses of the secondary outcomes (depressive symptom severity, anxiety symptom severity, quality of life utility scores, quality of life VAS, and days out of role) measured at six and 12 months. For these secondary outcomes, except days out of role, our outcome will be the change in scores between baseline and 6/12 months (whichever is relevant). For days out of role, we will use negative binomial regression where the outcome is the number of days totally and partially out of role at six or 12 months. All analyses will use multiple imputation.

Adherence-adjusted analysis

The effects of non-compliance on the estimated treatment effects will be investigated using a complier average casual effect (CACE) analysis [24]. Prior to data analysis, the study investigators

and data management team will review adherence with the intervention and construct three measures of adherence for each individual in the severe symptom intervention group (i.e., those allocated to care navigation). CACE analysis will therefore be performed for the primary outcome in the severe symptom group only.

The definitions below list the requirements that must be met for participants in the care navigation group to be considered to have received (adhered to) the intervention.

Definition of treatment adherence for the CACE analysis

1.	2.	3.
Participants attended at least one appointment with the care navigator AND There was a match between the patient priority and action plan	Participants attended at least one appointment with the care navigator AND There was a match between the patient priority and action plan AND There was a referral made to other services	Participants attended at least one appointment with the care navigator AND There was a match between the patient priority and action plan AND There was a referral made to other services AND The patient was provided with care package funding

We will conduct three separate CACE analyses. We will undertake this using two-stage least squares instrumental variable regression where the adherence variables are binary indicator variables capturing the definitions described the table above and trial arm used as the instrumental variable for adherence to treatment. Our analysis will control for baseline K10 scores and be estimated using multiple imputation.

Estimated treatment effect for primary and secondary outcomes

Estimated treatment effects for the primary and secondary outcomes will be presented as the difference in change in mean scores from baseline between the two trial arms (intervention-comparison) on their original metric. Days out of role will be presented as a rate ratio (i.e. on the exponential scale). The results for the primary analysis will also be presented as a standardised mean difference calculated relative to the pooled standard deviation (SD) from the baseline outcome scores. The estimated treatment effects with respective 95% confidence intervals (CI) and 2-sided p-value will be presented for the entire sample and by each symptom severity group as shown in Tables 2 and 3 for the six month outcome data. These tables will be replicated for the 12 month outcome data.

For the primary outcome, the estimated means for each trial arm with respective 95% confidence intervals will be plotted (*y*-axis) against time (baseline, 6 and 12 months; *x*-axis), for the entire sample and by each symptom severity group.

Sensitivity analyses for departures from the missing data assumption

For the primary analysis using multiple imputation (page 13), data are assumed to be missing at random (MAR), conditional on the variables included in the imputed model [25]. For the primary outcome (K10 score at six months) we will use a pattern mixture model to assess the robustness of the missing data assumption for the entire sample and for each symptom severity group.

Analysis for departures from MAR will be assessed by adding the quantity $\Delta = p_1\delta_1 - p_0\delta_0$ to the estimated treatment effect for the K10 score at six months, where p_i is the proportion of missing data at six months and δ_i the difference in mean K10 score between the participants with missing and those with observed responses in the intervention $(i = 1)$ and comparison $(i = 0)$ arms [26].

A range of values for δ_i will be considered for the difference in K10 scores between the participants with missing data and those observed at six months. Given that higher K10 scores indicate greater distress, negative values of δ_i assume that individuals with missing data have lower (better) K10 scores on average than observed individuals and positive values of δ_i assume that individuals with missing data have higher (worse) K10 scores than the observed mean score.

The primary analysis under MAR assumes that individuals with missing data have the same mean depressive symptom scores as those observed, that is $\delta_i = 0$ in both trial arms. For the sensitivity analyses, the difference between missing and observed K10 scores will be varied over the specified range of values for δ_i in the same way in both arms (that is, $\Delta = (p_1 - p_0)\delta$), vary in the intervention arm only and fixed at zero for the comparison arm $(\Delta = p_1\delta)$, and vary in the comparison arm and fixed at zero for the intervention arm $(\Delta = -p_0\delta)$.

The estimated treatment effect with respective 95% confidence intervals will be plotted on the *y*-axis in both trial arms, for selected parameter values of the difference between missing and observed mean score for K10 score (δ) at six months on the *x*-axis. A horizontal reference line will be plotted at zero on the *y*-axis, where positive values of the estimated treatment effect will indicate that the mean K10 score in the comparison arm is lower (better) than the intervention arm and negative values indicate that the intervention arm have lower (better) mean K10 than the comparison arm.

The analysis to assess robustness of missing data assumption may be repeated, as appropriate, for the secondary outcomes and by symptom severity group.

Economic evaluation

Perspective

The overall framework for the analysis of the economic evaluation will be a full economic evaluation using a within trial method as well as economic modelling to evaluate population level costs and effects. 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 patients when accessing health care. A partial societal perspective, which includes absenteeism and presenteeism effects on productivity for study participants, will be undertaken as a secondary analysis as recommended by the Second Panel on Cost-Effectiveness [27].

Whilst this evaluation will predominantly measure the resource use at a micro level (e.g., different types of services used), a gross costing approach will also be applied to the valuation of some items for pragmatic reasons (e.g., average hospital day cost). The gross costing approach where adopted should not affect the overall precision of the costing estimate since it is applied across both trial arms.

The reference year for the cost analyses will be 2018/2019.

Costs

The health sector costs include the cost to deliver the intervention as well as other health care services utilised by study participants during the study period. A partial societal perspective incorporates the cost of lost productivity. An impact inventory describing the outcomes and cost categories included within this economic analysis is provided in at Appendix A, as recommended by the Second Panel on Cost-Effectiveness in Health and Medicine [27].

The development costs of the DST have been previously estimated for use in the Australian population with depression (Target-D [28]). The current analysis will utilise the total development costs but will calculate an average cost estimate based on the Australian population with characteristics similar to the study entry criteria and likely to benefit from the intervention.

Additional costs for intervention delivery include the cost of iPads used to complete the DST in general practice waiting rooms. The cost of iPads will be provided by the study team and an average cost estimate will be calculated based on the likely lifespan of an iPad and the number of patients that can be screened within each PHN.

The intervention for the minimal/mild symptom group comprises services recommended through the DST and conveyed to patients via email. The use of additional services accessed will be captured through the self-completed RUQ, as explained in more detail below.

The intervention for the severe symptom group includes up to eight appointments with a care navigator that can provide access to additional funding ('care packages') to support service use. An average cost per care navigation session will be calculated by totalling care navigator costs (including salary and on-costs) and dividing by the total number of care navigation sessions. This care navigation session unit cost will then be applied to the actual number of sessions from participant records.

The cost of care packages provided to specific participants will be calculated from PHN records detailing requested services, actual use and cost paid by the PHN. A portion of the cost for some services could be claimed through MBS, however, this was not done due to administrative burden. The base case analysis will include the total cost of all services and a sensitivity analysis will be undertaken to understand the effect of claiming MBS reimbursement.

Health care and related care costs outside of the intervention

Additional health care and support services that participants, within both the intervention and comparison arms, access over the course of the trial period for the purpose of managing their mental health will be captured with the RUQ and additional administrative data where possible. The administrative data sets to be considered for use include: individual MBS, PBS, Primary Mental Health Care Minimum Data Set (PMHC MDS), and *headspace* data (for more information refer to the trial protocol [1]. All participants will be asked for additional separate consent to access the information held in these records for three months prior to trial participation and up to 12 months after enrolment.

Access to administrative data will only be available for those participants that provide this consent by completing a separate form. Participants may consent to provide access to all, some, or none of the requested datasets. Completed consent forms will be sent to the Australian Government Department of Human Services (MBS/PBS), the Department of Health (PMHC MDS), and *headspace* National Office at the end of the 12-month follow up period through a secure and confidential

process. Therefore, this information will not be available for initial analysis of six-month outcome data.

The analysis of six-month costs will utilise the information provided through the self-reported RUQ. The analysis of 12-month costs will also utilise the MBS/PBS and PMHC MDS data for participants who have provided consent to access this data. Preliminary analysis of the administrative data will be undertaken separately to the RUQ data.

PBS item prices will be used to calculate the government and patient out of pocket costs for covered medications reported in the RUQ [29]. Online Australian retail pharmacy sites will be accessed to determine patient costs for other medications and supplements not covered by the PBS [30]. Health professional visits will be costed using a weighted average cost paid by the government for the corresponding health professional, derived from the MBS item reports [31]. 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 for the base case analysis.

The reported number of times ambulance services were used by participants will be multiplied by an average ambulance service cost.

Emergency department visits will be costed using the average cost per emergency department presentation from the most recent Independent Hospital Pricing Authority (IHPA) data and inflated to our reference year if needed. The out of pocket cost for these services was reported by participants and will be added into the total out of pocket cost category.

Hospital stays will be costed using the IHPA National Weighted Activity Units (NWAUs) multiplied by the 2018/2019 National Efficient Price of $5,012 per National Weighted Activity Unit (IHPA). The specific Australian-refined diagnostic related groups (AR-DRGs) will be selected based on the reported reason for hospitalisation and the length of stay.

The partial societal perspective also incorporates effects on productivity. Participants were asked about the number of days (in the past six months) they have taken off from paid and unpaid work. They were also asked to report the number of days (in the past six months) 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 plus 25% overhead costs [32]. Time off from unpaid activities (i.e. housework) was valued at 25% of the average wage rate plus overhead costs to represent the value of participants' lost leisure time [33].

Presenteeism will be valued by multiplying the number of days reported working but bothered by mental health problems by 7.6 hours estimated in a full-time workday. The number of hours worked at reduced capacity will then be multiplied by 1 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 lost due to presenteeism which will then be valued using the average wage rate plus overhead costs.

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

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 with current evidence suggesting a threshold of around AU$28,000/QALY [34].The Australian value set for the EQ-5D-5L will be used to derive utility values at each assessment time point [35].The utility values at each time point will then be used to calculate total QALYs for each participant using the area under the curve method [36].

The K10 psychological distress score will also be utilised as an additional outcome measure in the economic analysis. This means that the difference in average total cost between arms will be compared to the average difference in the K10 score between arms as an alternate assessment of value for money also referred to as cost consequence analysis.

Statistical analysis of economic data

The statistical analyses for the economic evaluation will follow the principles detailed previously for the primary analysis and will employ an ITT approach, where all individuals randomised will be included in the analysis by their allocated trial arm status regardless of whether they received all, part or none of the intended treatments.

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 comparison arms at six months with adjustment for symptom severity group (minimal/mild vs. severe) and baseline K10 scores. Similar to the primary analysis, general practice site will not be included in the model for the primary analysis but will be included in sensitivity analysis (see next page). Separate GLMs will be used to estimate the difference in total societal costs and QALYs between arms at six months.

Negative binomial regression will be used to explore the between arm differences for the components of total cost including health care consultations, medications, ambulance, emergency department, hospital visits and lost productivity.

In addition to reporting descriptive statistics and differences between arms for costs and outcomes, incremental cost-effectiveness ratios (ICERs) will be calculated. ICERs will be calculated as the difference in cost (health sector and societal) divided by the difference in the effect or outcome between the two trial arms.

ICERs will be calculated as the difference in average costs between the two arms, divided by the difference in average outcome namely QALYs and K10 scores. Average ICERs and CIs will be calculated using a nonparametric bootstrap procedure, with 1,000 iterations to reflect sampling uncertainty. The bootstrapped ICERs and the CIs will be graphically represented on cost-effectiveness planes. A cost-effectiveness plane is a plot of the 1,000 bootstrapped incremental costs and outcomes across four quadrants. The north-east quadrant represents the intervention costing more as well as conferring greater benefits than the comparator. The south-east quadrant shows the proportion of iterations where the intervention costs less but incurs greater benefits than the comparator (i.e., a “dominant” intervention), the north-west quadrant shows the proportion of

iterations where the intervention incurs a cost but fewer benefits than the comparator (i.e., a "dominated" intervention) and, lastly, the south-west quadrant shows the proportion of iterations whereby the intervention costs less and has fewer benefits than the comparator group.

Sensitivity analyses

Sensitivity analyses will be conducted using complete-cases only using generalised linear models, with and without adjustment for general practice and other covariates. Complete cases will be records with resource use data observed at six months.

Additional sensitivity analyses will be undertaken to explore the effects of claiming MBS reimbursement for services provided through care packages; the effects of varying the cost of care navigation sessions to account for efficiencies that may be achieved over time; the opportunity costs of online therapies; and the effects of using the United Kingdom value set for the EQ-5D-5L [13].

Modelled economic evaluation

The costs and outcomes data from the within trial evaluation will then be used to evaluate the population cost-effectiveness of the intervention using economic modelling techniques. The modelling will incorporate two main extra components, the first being the costs of rolling out the Link-me interventions at an Australian population level – including both approaches to the minimal/mild and severe symptom groups. Secondly, the potential longer-term health benefits (and costs) at a population level will also be estimated. This will be undertaken using the epidemiological literature to estimate longer terms trajectories of severity states as well as resource use implications. More details regarding the modelled economic evaluation will be provided after the within trial economic evaluation has been completed. The modelling will only be undertaken if the within trial economic analysis finds that the Link-me intervention is cost-effective.

Results

The pages that follow present figure and table shells to be included in the Link-me report submitted to the Australian Government Department of Health and journal publications.

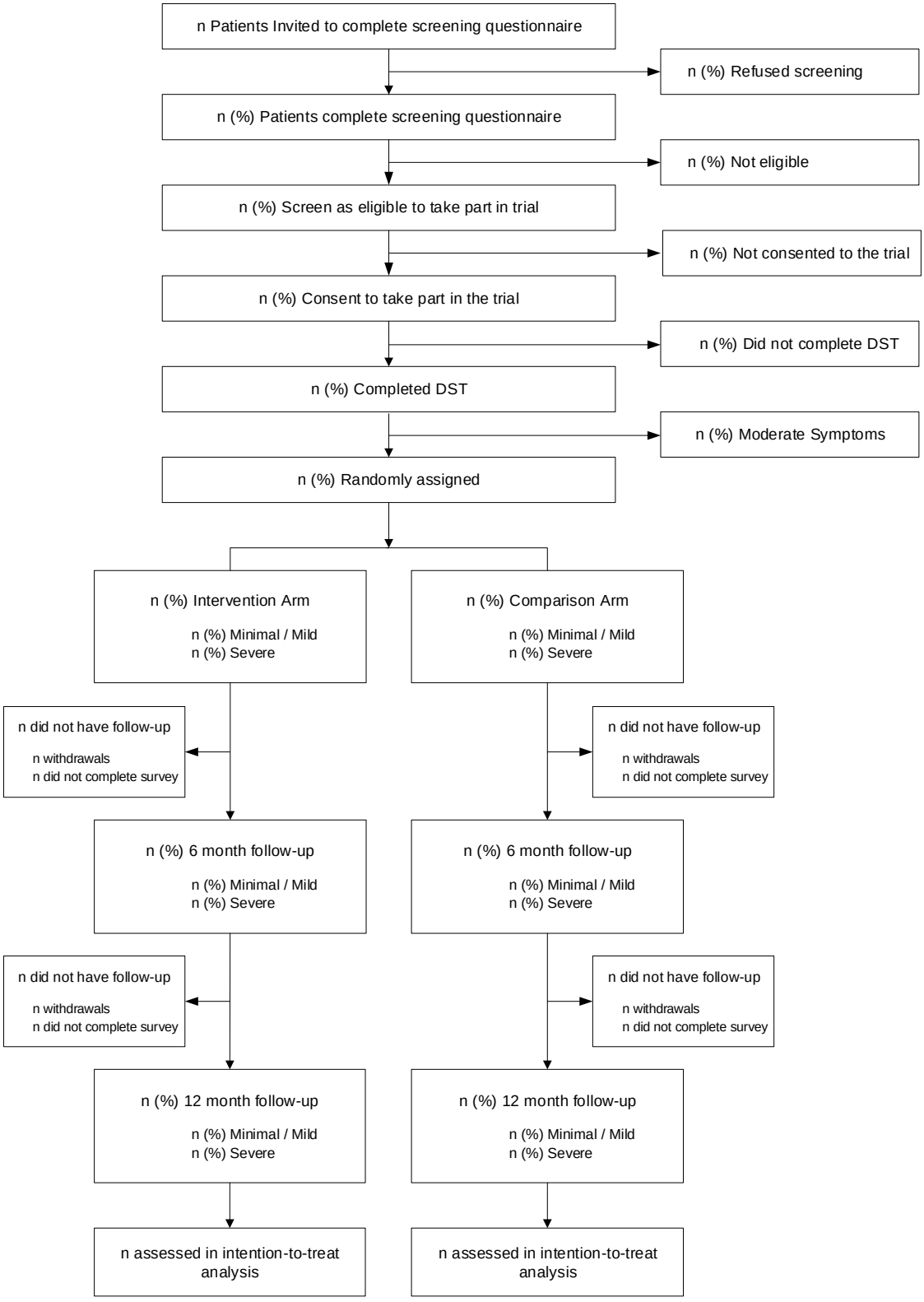


Figure 1. Trial profile

Note: Percentages at 6- and 12-month follow-up based on the total randomly allocated to each arm

Table 1. Baseline characteristics of participant according to trial arm, in total and stratified by symptom severity group

	All participants				Minimal/mild symptom group				Severe symptom group			
	Intervention arm (n=)		Comparison arm (n=)		Intervention arm (n=)		Comparison arm (n=)		Intervention arm (n=)		Comparison arm (n=)	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Age (years)												
Psychological distress (K10)												
Depressive symptom severity (PHQ 9)												
Anxiety symptom severity (GAD 7)												
Overall self-rated health (EQ-5D-5L)												
	Median	IQR	Median	IQR	Median	IQR	Median	IQR	Median	IQR	Median	IQR
Days totally out of role (K10+)												
Days partially out of role (K10+)												
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
Gender*												
Male												
Female												
Other												
Indigenous status												
Aboriginal												
Torres Strait Islander												
Aboriginal and Torres Strait Islander												
None of the above												
Language mainly spoken at home												
English												
Other												
Highest level of education completed												
Below Year 10												
Year 10 / equivalent												
Year 11 / equivalent												

	All participants		Minimal/mild symptom group		Severe symptom group	
	Intervention arm (n=)	**Comparison arm (n=)**	**Intervention arm (n=)**	**Comparison arm (n=)**	**Intervention arm (n=)**	**Comparison arm (n=)**
Year 12 / equivalent						
Certificate III/IV						
Advanced diploma / Diploma						
Bachelors degree						
Graduate diploma/Certificate						
Postgraduate degree						
Current employment						
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						
None of the above						
Main activity for those not working or looking for work						
Retired or voluntarily inactive						
Home duties						
Caring for children						
Studying						
Unable to work due to own illness, injury, or disability						
Caring for an ill or disabled person						
Working in an unpaid voluntary job						
Other						
Holds a health care card						
Managing on available income*						
Easily						
Not too bad						
Difficult some of the time						

	All participants		Minimal/mild symptom group		Severe symptom group	
	Intervention arm (n=)	**Comparison arm (n=)**	**Intervention arm (n=)**	**Comparison arm (n=)**	**Intervention arm (n=)**	**Comparison arm (n=)**
Difficult all of the time						
Impossible						
Live alone*						
Self-rated health						
Excellent						
Very good						
Good						
Fair						
Poor						
Long-term illness which limits daily activities*						
Reason for visit to GP						
Physical health						
Mental health and wellbeing						
Both physical and mental health						
None of these						
History of depression						
Currently taking medication for mental health						

Note: Sub-categories may be collapsed in final table published. *Item included in DST; IQR = Inter quartile range

Table 2. K10 psychological distress scores according to trial arm, in total sample and stratified by symptom severity group at 6 months

	All participants		Minimal/mild symptom group		Severe symptom group	
Intervention arm, n	n		n		n	
Comparison arm, n	n		n		n	
Mean change [1]						
Intervention arm, mean (SD)	mean (SD)		mean (SD)		mean (SD)	
Comparison arm, mean (SD)	mean (SD)		mean (SD)		mean (SD)	
Mean difference, Coef. (95% CI)						
Primary analysis [2]	estimate (95% CI)	p-value	estimate (95% CI)	p-value	estimate (95% CI)	p-value
Sensitivity analysis [3]	estimate (95% CI)	p-value	estimate (95% CI)	p-value	estimate (95% CI)	p-value
Sensitivity analysis [4]	estimate (95% CI)	p-value	estimate (95% CI)	p-value	estimate (95% CI)	p-value
CACE analysis [5]					estimate (95% CI)	p-value
CACE analysis [6]					estimate (95% CI)	p-value
CACE analysis [7]					estimate (95% CI)	p-value
Effect size, SMD (95% CI) [8]	estimate (95% CI)	p-value	estimate (95% CI)	p-value	estimate (95% CI)	p-value

Notes: SD = Standard deviation; Coef. = Estimated coefficient; CI = Confidence interval; d = Cohen's d. [1] Estimated using multiple imputation. [2] Mean for intervention arm minus mean for control arm estimated using linear regression adjusted for baseline outcome measure (all models) and symptom severity group (model with all participants only). Estimated using multiple imputation. [3] Sensitivity analysis using complete cases only with linear regression adjusted for baseline outcome measure (all models) and symptom severity group (model with all participants only). [4] Same as 3 but adjusted for general practice treated as random intercept. [5] CACE analysis: undertaken in the severe symptom severity group only. Conducted using two-stage least squares instrumental variable regression where the adherence variable is a binary coded variable representing participants attended at least one appointment with the care

navigator and there was a match between patient priorities and the action plan. Estimated using multiple imputation. [6] CACE analysis: Same as 5 except the adherence variable is a binary coded variable representing (1) participants attended at least one appointment with the care navigator and there was a match between patient priorities and the action plan and (2) a referral was made to other services. Estimated using multiple imputation. [7] CACE analysis: Same as 5 except the adherence variable is a binary coded variable representing (1) participants attended at least one appointment with the care navigator and there was a match between patient priorities and the action plan, (2) a referral was made to other services, and (3) the patient was provided with care package funding. Estimated using multiple imputation. [8] Mean difference in the primary analysis calculated relative to the pooled SD of baseline scores.

Table 3. Secondary outcomes according to trial arm, in total sample and stratified by symptom severity group at 6 months

	All participants	p-value	Minimal/mild symptom group	p-value	Severe symptom group	p-value
Depressive symptom severity (PHQ 9)						
Intervention arm[1]	mean (SD)		mean (SD)		mean (SD)	
Comparison arm[1]	mean (SD)		mean (SD)		mean (SD)	
Mean difference (95% CI)[2]	estimate (95% CI)	p-value	estimate (95% CI)	p-value	estimate (95% CI)	p-value
Anxiety symptom severity (GAD 7)						
Intervention arm[1]	mean (SD)		mean (SD)		mean (SD)	
Comparison arm[1]	mean (SD)		mean (SD)		mean (SD)	
Mean difference (95% CI)[2]	estimate (95% CI)	p-value	estimate (95% CI)	p-value	estimate (95% CI)	p-value
Quality of life -- Utility (EQ-5D-5L)						
Intervention arm[1]	mean (SD)		mean (SD)		mean (SD)	
Comparison arm[1]	mean (SD)		mean (SD)		mean (SD)	
Mean difference (95% CI)[2]	estimate (95% CI)	p-value	estimate (95% CI)	p-value	estimate (95% CI)	p-value
Quality of life -- VAS (EQ-5D-5L)						
Intervention arm[1]	mean (SD)		mean (SD)		mean (SD)	
Comparison arm[1]	mean (SD)		mean (SD)		mean (SD)	
Mean difference (95% CI)[2]	estimate (95% CI)	p-value	estimate (95% CI)	p-value	estimate (95% CI)	p-value
Days out of role						
Intervention arm[1]	mean (SD)		mean (SD)		mean (SD)	
Comparison arm[1]	mean (SD)		mean (SD)		mean (SD)	
Rate ratio, RR (95% CI)[3]	estimate (95% CI)	p-value	estimate (95% CI)	p-value	estimate (95% CI)	p-value

	All participants	p-value	Minimal/mild symptom group	p-value	Severe symptom group	p-value
Days limited in role						
Intervention arm[1]	mean (SD)		mean (SD)		mean (SD)	
Comparison arm[1]	mean (SD)		mean (SD)		mean (SD)	
Rate ratio, RR (95% CI)[3]	estimate (95% CI)	p-value	estimate (95% CI)	p-value	estimate (95% CI)	p-value

Notes: SD = Standard deviation; CI = Confidence interval. [1] Estimated using mutliple imputation. [2] Mean for intervention arm minus mean for comparison arm estimated using linear regression adjusted for baseline outcome measure (all models) and symptom group (model with all participants only). Estimated using multiple impution. [3] Ratio of the rate in the intervention arm divided by the rate in the comparison arm estimated using negative binomial regression adjusted for baseline outcome measure (all models) and symptom group (model with all participants only). Estimated using multiple imputation.

References

1. Fletcher S, Chondros P, Palmer VJ, Chatterton ML, Spittal MJ, Mihalopoulos C, et al. Link-me: Protocol for a randomised controlled trial of a systematic approach to stepped mental health care in primary care. Contemporary Clinical Trials 2019; 78: 63-75. doi: 10.1016/j.cct.2018.12.014.
2. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W. Process evaluation of complex interventions: Medical Research Council guidance. Br Med J 2015; 350. doi: 10.1136/bmj.h1258.
3. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: Validity of a two-item depression screener. Medical Care 2003; 41(11): 1284-92. doi: 10.1097/01.Mlr.0000093487.78664.3c.
4. Kroenke K, Spitzer RL, Williams JBW, Monahan PO, Lowe B. Anxiety disorders in primary care: Prevalence, impairment, comorbidity, and detection. Annals of Internal Medicine 2007; 146(5): 317-25.
5. Chondros P, Davidson S, Wolfe R, Gilchrist G, Dowrick C, Griffiths F, et al. Development and validation of a prognostic model for predicting depression severity in adult primary patients with depressive symptoms using the diamond longitudinal study. Journal of Affective Disorders 2018; 227: 854-60. doi: 10.1016/j.jad.2017.11.042.
6. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SLT, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychological Medicine 2002; 32(6): 959-76. doi: 10.1017/S0033291702006074.
7. Australian Bureau of Statistics. Use of the Kessler Psychological Distress Scale in ABS Health Surveys, Australia, 2007-08. 2012 [cited 2018 August]; Available from: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4817.0.55.001Chapter92007-08.
8. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a brief depression severity measure. Journal of General Internal Medicine 2001; 16(9): 606-13. doi: 10.1046/j.1525-1497.2001.016009606.x.
9. Kroenke K, Spitzer RL, Williams JBW, Löwe B. Psychiatry and primary care: The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: A systematic review. General Hospital Psychiatry 2010; 32: 345-59. doi: 10.1016/j.genhosppsych.2010.03.006.
10. Spitzer RL, Kroenke K, Williams JBW, Lowe B. A brief measure for assessing generalized anxiety disorder: The GAD-7. Archives of Internal Medicine 2006; 166(10): 1092-7. doi: 10.1001/archinte.166.10.1092.
11. Department of Health. Primary mental health care minimum data set: Scoring the Kessler-10 Plus. Canberra: Australian Government; 2018.
12. Herdman M, Gudex C, Lloyd A, Janssen MF, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of Life Research 2011; 20(10): 1727-36. doi: 10.1007/s11136-011-9903-x.
13. EuroQol. EQ-5D-5L user guide. 2015 [cited 2018 August]; Available from: https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-5L_UserGuide_2015.pdf.
14. Norman R, Cronin P, Viney R. A pilot discrete choice experiment to explore preferences for EQ-5D-5L health states. Appl Health Econ Health Policy 2013; 11(3): 287-98. doi: 10.1007/s40258-013-0035-z.
15. Pink B. Socio-Economic Indexes for Areas (SEIFA), 2011. Canberra: Australian Bureau of Statistics; 2013.
16. StataCorp. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC; 2017.
17. National Health and Medical Research Council. National statement on ethical conduct in human research 2007 (Updated 2018). Canberra: Commonwealth of Australia; 2007.

18. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2019.
19. White IR, Carpenter J, Horton NJ. Including all individuals is not enough: Lessons for intention-to-treat analysis. Clinical Trials 2012; 9(4): 396-407. doi: 10.1177/1740774512450098.
20. Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: Current practice and problems. Statistics in Medicine 2002; 21(19): 2917-30.
21. Piantadosi S. Clinical trials: A methodological perspective. Hoboken, NJ: John Wiley & Sons; 2017.
22. Lee KJ, Carlin JB. Multiple imputation for missing data: fully conditional specification versus multivariate normal imputation. American Journal of Epidemiology 2010; 171(5): 624-32. doi: 10.1093/aje/kwp425.
23. Rubin DB. Multiple imputation for nonresponse in surveys. New York: John Wiley & Sons, Inc.; 1987.
24. 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. doi: 10.1191/0962280205sm403oa.
25. White IR, Kalaitzaki E, Thompson SG. Allowing for missing outcome data and incomplete uptake of randomised interventions, with application to an Internet-based alcohol trial. Statistics in Medicine 2011; 30(27): 3192-207. doi: 10.1002/sim.4360.
26. Little RJA. Pattern-mixture models for multivariate incomplete data. Journal of the American Statistical Association 1993; 88(421): 125-34. doi: 10.1080/01621459.1993.10594302.
27. Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, et al. Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: Second Panel on Cost-Effectiveness in Health and Medicine. Journal of the American Medical Association 2016; 316(10): 1093-103.
28. Gunn J, Wachtler C, Fletcher S, Davidson S, Mihalopoulos C, Palmer V, et al. Target-D: A stratified individually randomized controlled trial of the diamond clinical prediction tool to triage and target treatment for depressive symptoms in general practice: Study protocol for a randomized controlled trial. Trials 2017; 18: 342. doi: 10.1186/s13063-017-2089-y.
29. Pharmaceutical Benefits Scheme. Ex-manufacturer price. 2019 [cited 2019 May]; Available from: http://www.pbs.gov.au/info/industry/pricing/ex-manufacturer-price.
30. Chemist Warehouse. Available from: https://www.chemistwarehouse.com.au/.
31. Department of Human Services. Medicare item reports. 2019 [cited 2019 May]; Available from: http://medicarestatistics.humanservices.gov.au/statistics/mbs_item.jsp.
32. Australian Bureau of Statistics. Average weekly earnings, Australia, Nov 2018. Canberra: ABS; 2019 [cited 2019 May]; Available from: https://www.abs.gov.au/ausstats/abs@.nsf/mf/6302.0.
33. 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.
34. Edney LC, Haji Ali Afzali H, Cheng TC, Karnon J. Estimating the reference incremental cost-effectiveness ratio for the Australian health system. Pharmacoeconomics 2018; 36(2): 239-52. doi: 10.1007/s40273-017-0585-2.
35. Viney R, Norman R, Brazier J, Cronin P, King MT, Ratcliffe J, et al. An Australian discrete choice experiment to value EQ-5D health states. Health Economics 2014; 23(6): 729-42. doi: 10.1002/hec.2953.
36. Glick HA, Doshi JA, Sonnad SS, Polsky D. Economic evaluation in clinical trials (2nd ed.). Oxford: Oxford University Press; 2014.

Appendix A

Impact Inventory as recommended by the Second Panel on Cost-Effectiveness in Health and Medicine

Sector	Type of Impact	Included in this reference case analysis from …perspective?		Notes on sources of evidence
		Health Care Sector	Societal	
Formal Health Care Sector				
Health	Health outcomes (effects)			
	Longevity effects			
	Health -related quality-of-life effects	✓	✓	EQ-5D-5L
	Other health effects (eg, adverse events and secondary transmissions of infections)	✓	✓	K10 score
	Medical costs			
	Paid for by third-party payers	✓	✓	Medications, consultations, emergency department, hospital care reimbursed by government
	Paid for by patients out-of-pocket	✓	✓	Medications, consultations, hospital care, emergency department fees paid by participants
	Future related medical costs (payers and patients)			

	Future unrelated medical costs (payers and patients)			
Informal Health Care Sector				
Health	Patient-time costs	NA		
	Unpaid caregiver-time costs	NA		
	Transportation costs	NA		
Non-Health Care Sectors (with examples of possible items)				
Productivity	Labour market earnings lost	NA	✓	
	Cost of unpaid lost productivity due to illness	NA	✓	
	cost of uncompensated household production	NA		
Consumption	Future consumption unrelated to health	NA		
Social Services	Cost of social services as part of intervention	NA		
Legal or criminal justice	Number of crimes related to intervention	NA		
	Cost of crimes related to intervention	NA		
Education	Impact of intervention on educational achievement of population	NA		
Housing	Cost of intervention on home improvements (e.g., removing lead paint)	NA		
Environment	production of toxic waste pollution by intervention	NA		
Other (specify)		NA		

Template based on Figure 1 from Sanders et al [27].