ANTidepressants to prevent reLapse in dEpRession (ANTLER)

Statistical analysis plan

Written by Louise Marston (LM), Nick Freemantle (NF), Caroline Clarke (CC), Rachael Hunter (RH), Larisa Duffy (LD) and Glyn Lewis (GL)
Last updated on 16/01/2020 by Louise Marston
Updated on 18/12/2019 by Louise Marston
Updated on 11/12/2019 by Louise Marston
Updated on 06/12/2019 by Louise Marston
Updated on 21/11/2019 by Louise Marston
Updated on 03/10/2019 by Louise Marston
Updated on 19/08/2019 by Louise Marston
Updated on 30/05/2019 by Louise Marston
Updated on 07/03/2019 by Louise Marston
Updated on 06/03/2019 by Louise Marston and Nick Freemantle
Updated on 13/11/2018 by Louise Marston
Updated on 23/10/2018 by Louise Marston
Updated on 12/10/2018 by Larisa Duffy
Updated on 21/09/2018 by Louise Marston
Updated on 11/09/2018 by Louise Marston and Larisa Duffy
Updated on 10/09/2018 by Louise Marston
Updated on 19/07/2018 by Caroline Clarke and Rachael Hunter
Updated on 09/02/2018 by Louise Marston
Updated on 11/01/2018 by Louise Marston
Updated on 30/10/2017 by Louise Marston
Updated on 26/10/2017 by Louise Marston
Version: 3.0

Introduction

This analysis plan sets out the methods of analysing the predetermined primary, secondary and health economic outcomes for ANTLER, which will be reported in the National Institute for Health Research, Health Technology Assessment report at the end of the trial and also in the main peer review paper(s) to result from this randomised controlled trial.

The analysis and reporting of this trial will conform to the CONSORT and CHEERS statements[1-4] and the appropriate standard operating procedures written by Priment Clinical Trials Unit.

Further information on this trial can be found in the protocol version 7.0 (26/11/2018). The protocol is stored on: S:\Pop_Health\PCPH_Priment\Projects\Current\CTIMPS\ANTLER\6. Protocol\ANTLER_protocol_v7_20181126.pdf.

Trial summary

Aim

To estimate the effectiveness and cost-effectiveness of antidepressant medication in preventing relapse in UK primary care in people who have had two or more episodes of depression (including the current episode), have taken antidepressants for at least nine months and are now well enough to consider stopping the antidepressant.

Objectives

To estimate the difference in time to depressive episode between randomised groups.

To estimate the difference in depression and anxiety symptoms between randomised groups.

To estimate the difference in adverse effects of antidepressants by randomised groups.

To determine the difference in withdrawal symptoms between randomised groups.

To estimate the difference in health related quality of life between randomised groups.

To compare the relative cost-effectiveness of the two arms of the trial.

Study population

Inclusion criteria

Participants will be included if they:

- have had at least two episodes of depression
- are aged 18-74 years
- have been taking antidepressants for 9 months or more and are currently taking: citalopram 20mg, sertraline 100mg, fluoxetine 20mg or mirtazapine 30mg
- have satisfactory adherence to medication using a 5-item self-report measure of compliance[5]. Given the relatively long half-life of antidepressant medication, individuals who have forgotten to take one or two tablets will not be excluded and this will be established with an extra question "Did you forget to take 2 days of your medication in a row?" Therefore, our criteria defined people as adherent if (1) they scored zero on all four questions (2) they scored one and said No to the extra question (3) scored 2 because of 'forget' and 'careless' questions and said No to extra question.
- are considering stopping their antidepressant medication.

Exclusion criteria

Participants will be excluded if they:

- meet internationally agreed (ICD10) criteria for a depressive illness
- have bipolar disorder, psychotic illness, dementia, alcohol or substance dependence or a terminal illness
- are not able to complete self-administered questionnaires in English
- have contraindications for any of the prescribed medication
- are pregnant or intend to get pregnant within the next 12 months
- are using monoamine oxidase inhibitors
- have allergies to placebo excipients
- are currently enrolled in another Clinical Trial of an Investigational Medicinal Product

Trial design

The study is a double blind individually randomised parallel group controlled trial.

Randomised treatments

At baseline participants will be taking either citalopram 20mg, sertraline 100mg, fluoxetine 20mg or mirtazapine 30mg.

Intervention:

One month of the same medication as at baseline at half the dose (citalopram 10mg, sertraline 50mg, or mirtazapine 15mg), followed by a quarter of the dose for a month taken as half the dose and placebo on alternate days and then taking placebo for the remainder of the study. There is no half dose capsule for fluoxetine so those taking fluoxetine at baseline who are allocated to the placebo arm will have one month of fluoxetine at half the dose, taken by alternating between a 20mg tablet and placebo tablet for one month. As no quarter dose of fluoxetine is available, those taking fluoxetine who are allocated to the placebo arm of the trial will take a placebo from the second month of the study.

Control
This is to remain on their current medication (which they were taking at baseline).

Sample size
The Geddes[6] systematic review estimated a reduction in odds of relapse of 70%, Kaymaz[7] 65%, Glue[8] 65% and NICE[9] about 50%. Between 15% and 22% of those on active drug relapsed in 12 months. To detect the difference between relapse rates of 15% (continuation) and 30% (withdrawal) (hazard ratio 0.46), or 20% (continuation) and 35% (withdrawal) (hazard ratio 0.52) will require sample sizes of respectively 333 and 383 for 90% power at the 5% significance level. Allowing for 20% attrition we therefore propose to recruit 479 participants. In the COBALT trial[5] 84% completed assessments at 12 months follow up so we will work to achieve this figure or higher.

Randomisation
Randomisation will be concealed by using a remote computerised system provided by Sealed Envelope[10] and the pharmacy will be directly informed of the randomisation outcome. The randomisation will be minimised by the four study centres, the four medications and severity of depressive symptoms at baseline (two categories).

Blinding
The data will be analysed blind to allocation by the statisticians and the health economists. One statistician (LM) may become unblinded during the course of their work with the Data Monitoring Committee (DMC). If this occurs, they will no longer attend the Trial Management Group meetings to prevent further unblinding of the trial team.

Outcomes

Primary outcome
Time in weeks to the beginning of the first episode of depression after randomisation. This will be measured using the Clinical Interview Schedule – Revised (CISR)[11], that will ask about the previous 12 weeks at all follow up points. Only the five questions (depression, depressive ideas, concentration, sleep and fatigue) used for a depression diagnosis will be asked along with questions asking about symptoms. Further questions will be asked to determine the time to the nearest week when the score was two or more. An episode of depression will be defined as scoring two or more on the depression section of the CISR for at least two weeks. The participant will need to score on at least one of the two mandatory depression questions:
"Almost everyone becomes low in mood or depressed at times. Has there been a time in the past three months when you had a spell of feeling sad, miserable or depressed?" or "In the past three months, have you been able to enjoy or take an interest in things as much as you usually do?"

If the participant answers Yes to the first question and/ or No to the second question, then they will be asked some duration questions and they will need to indicate that they have experienced low mood and/ or lack of enjoyment for at least two weeks in order to count it as an episode. In addition to the mandatory questions, the participants will need to score at least 1 on any of the additional depressive symptoms during the worst week in the past three months: fatigue, sleep problems, restlessness, suicidal thoughts, hopelessness, feeling low for prolonged periods, unresponsiveness of mood, retardation, loss of sexual interest, lack of concentration, reduced self-esteem and feeling of guilt.

The beginning of that episode (and date used for this outcome), will be the first week the participant scores two or more on the CISR depression scale.

Secondary outcomes
Secondary outcomes will be at 3, 6, 9 and 12 months separately.

Depressive symptoms (PHQ9)[12, 13] This is a nine item questionnaire. Each item has four responses ranging from not at all (0) to nearly every day (3). The score from each item is added to give a total ranging from 0 to 27. If there are one or two items missing from a participant's questionnaire, items will be replaced by the mean of the items present. If there are more than two items missing, the questionnaire will be considered missing for that participant.

Anxiety symptoms (GAD7)[14]. This is a seven item questionnaire. Each item has four possible responses ranging from not at all (0) to nearly every day (3). The score from each item is added to give a total ranging from 0 to 21. If there are one or two items missing from a participant's questionnaire, items will be replaced by the mean of the items present. If there are more than two items missing, the questionnaire will be considered missing for that participant.

Adverse effects of antidepressants using a modified Toronto side effects scale[15, 16]. This is a 13 item measure for males and females and an open-ended item to report any other side effects. For each item it asks how often the side effect has been present in the past two weeks; none, several days, more than half the days or nearly every day. Scores from each item will be added to give an overall score between 13 and 52.

Health related quality of life using the SF12[17, 18]. The physical and mental component scores will be analysed separately.

Withdrawal symptoms – based on Rosenbaum[19], with 15 items relating to the commonest symptoms enquired about. Scores are calculated by adding the number of "new symptom" and "old symptom but worse" that are reported ranging from 0 to 15.

Time to stopping study medication The exact date at which the medication is stopped will be recorded for those who stop early. For those who complete the course of medication the date will be the date of the last interview, or the date they took the last dose of study medication, whichever is later.

Quality-adjusted life-years (QALYs) using the EQ-5D-5L[20]. The QALYs for each arm will be calculated as the area under the line connecting utility scores calculated from EQ-5D-5L responses collected at each time point (baseline, 3, 6, 9, 12 months) and adjusting for baseline values.

Global Rating Question Participants were asked at baseline, 3, 6, 9 and 12 months. "Compared to when we last saw you, how have your moods and feelings changed?" Responses were: 'I feel a lot better'; 'I feel slightly better'; 'I feel about the same'; 'I feel slightly worse'; 'I feel a lot worse'. We will create a dichotomous variable: feeling worse (1) and feeling the same or better (0).

Primary healthcare resource use collected from GP electronic records

Client Service Receipt Inventory (modified) for other health and social care resource use and wider societal impact.

Mechanistic outcomes

The mechanistic outcomes will not be reported in the HTA report or in the main trial paper that reports the clinical outcomes as they were not funded by the HTA. They will be reported in separate subsequent paper(s).

Data collection

Baseline

Revised clinical interview schedule CISR
Past medical history questions including any physical illness contraindications and past psychiatric treatments
Sociodemographic and other background information
Depressive symptoms (PHQ9)
Anxiety symptoms (GAD7)
EQ-5D-5L
Adverse effects of antidepressants
Questions about other medications
Health related quality of life (SF12)
Health and social care resource use and wider societal costs
Withdrawal symptoms
Word recall task
Go–no go task
Faces task
Global rating question

6 weeks

Depressive symptoms (PHQ9)
Anxiety symptoms (GAD7)
Adherence to study medication
Questions about other medications
Health related quality of life
Withdrawal symptoms
Guess whether on active drug or placebo
Global rating question

3 months

Depressive symptoms (PHQ9)
Anxiety symptoms (GAD7)
Revised clinical interview schedule CISR
EQ-5D-5L
Adverse effects
Adherence to study medication
Questions about other medications
Health related quality of life (SF12)
Withdrawal symptoms
Guess whether on active drug or placebo
Word recall task
Go–no go task
Faces task
Global rating question

6 months

Depressive symptoms (PHQ9)
Anxiety symptoms (GAD7)
Revised clinical interview schedule CISR
EQ-5D-5L
Adverse effects
Adherence to study medication
Questions about other medications
Health related quality of life (SF12)
Health and social care resource use and wider societal costs

Withdrawal symptoms
Guess whether on active drug or placebo
Global rating question

9 months
Depressive symptoms (PHQ9)
Anxiety symptoms (GAD7)
Revised clinical interview schedule CISR
EQ-5D-5L
Adverse effects
Adherence to study medication
Questions about other medications
Health related quality of life (SF12)
Withdrawal symptoms
Guess whether on active drug or placebo
Global rating question

12 months
Depressive symptoms (PHQ9)
Anxiety symptoms (GAD7)
Revised clinical interview schedule CISR
EQ-5D-5L
Adverse effects
Adherence to study medication
Questions about other medications
Health related quality of life (SF12)
Health and social care resource use and wider societal costs
GP appointments and medication
Withdrawal symptoms
Guess whether on active drug or placebo
Word recall task
Go–no go task
Faces task
Global rating question

Data entry
Data will be entered using a web based system set up by Sealed Envelope[10]. This has been set up so that, it mirrors the data collection sheets in order. It also has range checks, consistency checks and for closed questions gives a number of options plus "other" where appropriate. Researchers who will be entering the data will have no access to the group allocation through this system. Data will be checked by a statistician and health economist before analysis and any problems reported to the Trial Manager, who will rectify them as appropriate before database locking and data analysis.

Statistical analyses
Secondary outcomes will be analysed using Stata v16[26], and the primary outcome will also be analysed using SAS v9.4 or above[27].

Interim analyses
There are no planned analyses. However, this does not preclude the DMC from requesting interim analyses.

Final analyses
All analyses will be complete case, intention to treat (defined as all patients randomised, analysed according to their randomised group regardless of treatment received).

The CONSORT[1-3] flow diagram will be constructed by/ in collaboration with the Trial Manager who will have logs of patients who do and do not agree to take part in the study. It will include number of patients randomised to each arm of the trial, and the numbers who have follow up data available.

Descriptive statistics
Initial analyses will look at summary statistics for all variables, both overall and by randomised group. Summary statistics for continuous variables will be mean, median, SD, lower quartile, upper quartile and reported appropriately according to distribution.

Analysis of the primary outcome
The primary outcome will be analysed using an exact Cox proportional hazards model (to account for ties), accounting for the depressive symptom score from the CIS-R at baseline.

Analysis of the secondary outcomes
Continuous secondary outcomes will be analysed using mixed effects linear regression with two observations per participant, including the baseline value. Variables indicating time and randomised group will be included in the models as fixed effects, and participants grouped using random intercept terms.

We will conduct a supportive analysis using all observations for a participant in a further mixed effects regression model for each continuous outcome.

Time to stopping study medication will be analysed using Cox proportional hazards modelling.

The global rating question will be analysed using mixed effects logistic regression in a similar way to the continuous secondary outcomes.

We will examine the test retest reliability of the PHQ9, GAD7, Retrospective CIS-R, adverse effects and withdrawal symptoms. Participants will be asked to repeat those questionnaires at one of the follow up appointments.

Missing data
Baseline predictors of missingness of the primary outcome will be examined. Multiple imputation will not be carried out because the assumption of missing at random is unlikely to apply.

Supportive analyses
There will be several supportive analyses for the primary outcome.

1 We will examine the validity of the underlying assumption of the primary analysis for constant proportional hazards, examining survival curves and fitting a time dependent covariate model to examine linearity with time. Where a significant departure from constancy exists, we will utilise a log rank test to provide the principal analysis.

2 Including the minimization variables (centre, antidepressant medication, level of depression) as participant level explanatory variables (factors).

3 Various assumptions about the reasons for missing data will be investigated. This will include imputing a good outcome for those in the control group who have missing outcome data and a poor outcome for those in the intervention group who have missing outcome data. For secondary outcomes, we will adjust for variables at baseline associated with missing data.

4 We will also consider undertaking an analysis where we condition the effects on the propensity of participants to adhere to the randomised treatment on the basis of their baseline characteristics.

5 We will examine the possible effects of lack of adherence to study treatments using a mediation analysis, including adherence to study medication as a mediator[28]

Subgroup analyses

Subgroup analyses will be carried out on time to relapse (primary outcome), PHQ9, GAD7 and global rating question (dichotomous). Analyses for the PHQ9, GAD7 and global rating question will be carried out at each time point. These will be carried out using interactions between the randomisation variable and the variables of interest detailed below. Other variables in these models will be as for the primary analyses of the primary and secondary outcomes.

- Antidepressant medication (dropping Mirtazapine due to small numbers)
- Residual CISR depression score (depression, depressive ideas, fatigue, concentration and sleep)
- CISR anxiety score (worry, anxiety, phobias, panic, worry about physical health) scores at baseline
- Two or more previous episodes of depression at baseline versus 0 or 1 previous episodes
- Age when first aware of problems with depression

Health economic analysis

Aim

We will calculate the mean incremental cost per quality adjusted life year (QALY) gained of replacing antidepressant with placebo compared to antidepressant maintenance over 12 months, from an NHS and social care perspective, using trial data.

The main outcome measures to be used in this analysis are health care resource use costs relating to treatment for mental health issues, and health-related quality of life.

Quality of life

Quality of life outcome measures collected during the trial are the SF-12[17, 18] and EQ-5D-5L[20]. The primary economic analysis will use patients' responses to the EQ-5D-5L to calculate quality-adjusted life-years (QALYs). QALYs represent both the quality and quantity of life, quality being measured by utility scores derived from patient-reported quality of life outcome instruments. QALYs are calculated by multiplying the utility weight for a health state by the length of time spent in that health state, and summing across health states over the time period of interest. A utility score of 1 represents perfect health and a utility of 0 represents death; negative values, representing states worse than death, are possible. QALYs are the recommended outcome for use in economic evaluations in the UK as they are a common unit that allows for comparable decisions about resource allocation across different health conditions. In the UK, NICE recommends that QALYs are calculated using utility scores generated by the EQ-5D, a 5-item, self-rated instrument covering mobility, self-care, usual activities, pain/ discomfort and anxiety/ depression[29]. The primary cost-utility analysis will use the NICE-recommended mapping algorithm to obtain utility weights[30].

The mean difference in the area under the curve for each group will be calculated using the available data as collected at baseline, 3, 6, 9 and 12 months (see Data Collection section), adjusting for baseline EQ-5D-5L score. The regression analysis will include covariates for randomisation to calculate treatment effect and depressive symptom score at baseline. Trial-period QALYs will then be calculated using the estimated utility weights. We will assess the

impact of missing data, and identify any baseline patient characteristics that are related to missing QALYs. Predictors of missing QALYs will be explored for use as covariates in the regression analysis and multivariate imputation by chained equations (MICE) will be used if preferable. For patients who die during follow-up, we will assume a straight line from the last available utility value to a zero at the date of death. Bootstrapping will be used to calculate 95% confidence intervals.

Costs

We will report the mean cost per patient in each arm, making similar baseline adjustments as discussed above for QALYs. This will be calculated using health care resource use data collected from trial participants via the Health and Social Care Resource Use questionnaire referred to above, which is a modified version of the Client Service Receipt Inventory (CSRI), at baseline, 6, and 12 months, requesting details of the previous 6 months' resource use each time. This will include information on community and acute care health service contacts, mental health community and inpatient service use, social care, employment and welfare payments. Data on primary care appointments and medications will be collected by a research assistant from participants' GP notes. Services will be costed using nationally published sources[31], and medications using the British National Formulary. The cost of antidepressant maintenance will be calculated for the control group. For the primary analysis costs will be from the NHS and social care perspective. A secondary analysis from the societal cost perspective will also be conducted, including data on employment collected from trial participants.

Difference in costs between the two groups will be calculated using ordinary least squares regression analysis and bootstrapping to account for the skewed nature of the data. Randomisation, baseline costs and depressive symptom score at baseline will be included as covariates in the regression analysis, with the beta coefficient for the randomisation coefficient reported with 95% confidence intervals.

Outputs from the analysis will be:

- Descriptive statistics of EQ-5D-5L and associated algorithm[32] to calculate the utility score
- Mean total patient-level QALYs for each arm, using utility scores derived from the EQ-5D-5L using the mapping algorithm recommended by NICE[30]
- Mean cost per patient of maintenance treatment (control arm) and reduction in dosage and replacement with placebo (intervention arm), broken down according to medication costs and other costs
- Mean total health and social care cost per patient in each arm
- Mean incremental cost per QALY of replacing antidepressant use with placebo compared to maintenance on antidepressant
- 95% confidence intervals generated using bootstrapping
- Cost-effectiveness acceptability curve generated using bootstrapped results
- Sensitivity analysis of the above results using utility scores derived instead from the new EQ-5D-5L value set for England[32].
- Sensitivity analysis of the above results using utility scores derived instead from the SF-6D scoring system, based on the SF-12 responses.

Incremental cost-effectiveness ratio (ICER)

The headline statistic which will be calculated using the costs and QALYs described above is the incremental cost-effectiveness ratio (ICER). This is the cost per unit change in QALYs on replacing antidepressant with placebo compared to maintenance antidepressant treatment in the trial population.

$$ICER = \frac{cost\ in\ placebo\ arm - cost\ in\ maintenance\ arm}{QALYs\ in\ placebo\ arm - QALYs\ in\ maintenance\ arm} \qquad \text{Equation 1}$$

In this analysis, we will calculate the QALYs per patient in each arm with adjustment for baseline values, and including the covariates stated above, using linear regression, meaning that the value to go into **Error! Reference source not found.** is the treatment effect coefficient from Equation 1:

$$Q_i = \beta_0 + \beta_1(j_i) + \beta_2(u_{i,t=0}) \qquad \text{Equation 1}$$

where Q_i = total QALYs for individual *i*; *j* is the treatment group (*j*=0 for maintenance and *j*=1 for placebo arm) β_1 = treatment effect coefficient; β_2 = baseline utility score coefficient[33].

Sensitivity analysis

We will conduct one and two-way sensitivity analyses for any assumptions made and subgroup analyses as identified. Missing data will be handled in the same way as the main statistical analysis, including examination of predictors of missingness, and with the primary analysis being an intention-to-treat analysis and secondary analyses taking into account assumptions about missingness and multiple imputation.

Cost-effectiveness acceptability curve

Bootstrapping will be used to construct confidence intervals and the results of the non-parametric bootstrapping will be presented on a cost-effectiveness plane, and a cost-effectiveness acceptability curve of the probability that replacing antidepressant with placebo is cost-effective compared to antidepressant maintenance for a range of values of willingness to pay for a QALY gained will be reported.

References

1 Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomised controlled trials. The CONSORT statement. JAMA 1996;276:637-9.

2 Moher D, Schulz KF, Altman DG, CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 2001;357:1191-4.

3 Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c332

4 Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, Augustovski F, Briggs AH, Mauskopf J, Loder E Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement BMJ. 2013;346:f1049

5 Wiles N, Thomas L, Abel A, Ridgway N, Turner N, Campbell J, et al. Cognitive behavioural therapy as an adjunct to pharmacotherapy for primary care based patients with treatment resistant depression: results of the CoBalT randomised controlled trial. Lancet. 2013;381:375-84

6 Geddes JR, Carney SM, Davies C, Furukawa TA, Kupfer DJ, Frank E et al. Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review. Lancet. 2003;361(9358):653-61

7 Kaymaz N, van Os J, Loonen AJ, Nolen WA. Evidence that patients with single versus recurrent depressive episodes are differentially sensitive to treatment discontinuation: a meta-analysis of placebo-controlled randomized trials. J Clin Psychiatry. 2008;69(9):1423-36

8 Glue P, Donovan MR, Kolluri S, Emir B. Meta-analysis of relapse prevention antidepressant trials in depressive disorders. The Australian and New Zealand journal of psychiatry. 2010;44(8):697-705.

9 NICE. The NICE Guideline on the Treatment and Management of Depression in Adults Updated Edition. In: Excellence NIfHaC, editor. London 2010.

10 http://www.sealedenvelope.com/ accessed July 2018

11 Lewis G, Pelosi AJ, Araya R, Dunn G. Measuring psychiatric disorder in the community: a standardised assessment for use by lay interviewers. Psychological Medicine. 1992;22:465-86.

12 Gilbody S, Richards D, Brealey S, Hewitt C. Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis. J Gen Intern Med. 2007;22(11):1596-602

13 Kroenke K, Spitzer RL, Williams JB 2001 The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9): 606-613

14 Spitzer RL, Kroenke K, Williams JBW, Löwe B A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7 2006; 166(10) 1092-1097

15 Vanderkooy J, Kennedy SH, Bagby RM (2002) Antidepressant side effects in depression patients treated in a naturalistic setting: a study of bupropion, moclobemide, paroxetine, sertraline, and venlafaxine. Can J Psychiatry 47:174–180

16 Crawford AA, Lewis S, Nutt D, Peters TJ, Cowen P, O'Donovan MC, et al. Adverse effects from antidepressant treatment: randomised controlled trial of 601 depressed individuals. Psychopharmacology (Berl). 2014;231(15):2921-31

17 Ware J Jr, Kosinski M, Keller SD (1996). A 12-Item short-form health survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care, 34*(3), 220–233

18 Ware JE, Kosinski M, Turner-Bowker DM, Gandek B SF-12v2: How to score version 2 of the SF-12 health survey (with a supplement documenting version 1); Lincoln, RI: QualityMetric Inc, 2002

19 Rosenbaum JF, Fava M, Hoog SL, Ascroft RC, Krebs WB Selective serotonin reuptake inhibitor discontinuation syndrome: a randomized clinical trial. Biol Psychiatry.1998;44(2):77-87.

20 Herdman M, Gudex C, Lloyd A, Janssen M, 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: an international journal of quality of life aspects of treatment, care and rehabilitation. 2011;20(10):1727-36.

21 Harmer CJ, Goodwin GM Cowen PJ Why do antidepressants take so long to work? A cognitive neuropsychological model of antidepressant drug action The British Journal of Psychiatry 2009;195:102–108

22 Guitart-Masip M, Economides M, Huys QJM, Frank MJ, Chowdhury R, Duzel E, et al. Differential, but not opponent, effects of l-DOPA and citalopram on action learning with reward and punishment. Psychopharmacology (Berl). 2014;231:955–966.

23 Button K, Lewis G, Penton-Voak I, Munafò M Social anxiety is associated with general but not specific biases in emotion recognition Psychiatry Research 2013;210(1):199-207

24 Lundqvist D, Flykt A, O¨ hman A. The Karolinska Directed Emotional Faces – KDEF. Karolinska Institutet, 1998

25 Tiddeman BP, Burt M, Perrett DI. Prototyping and transforming facial textures for perception research. IEEE Comput Graph Appl 2001; 21: 42–50.

26 StataCorp. 2019. *Stata Statistical Software: Release 16*. College Station, TX: StataCorp LLC

27 SAS software, Version x of the SAS System for Windows. Copyright © [year of copyright] SAS Institute

28 Emsley R, Dunn G, White IR Mediation and moderation of treatment effects in randomised controlled trials of complex interventions Statistical Methods in Medical Research 2010; **19**: 237–270

29 NICE. Guide to the methods of technology appraisal. [Online].; 2013 [cited 2015 January. Available from: http://publications.nice.org.uk/pmg9.

30 van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, Lloyd A, Scalone L, Kind P, Pickard AS Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets; Value Health. 2012 Jul-Aug;15(5):708-15. doi: 10.1016/j.jval.2012.02.008. Epub 2012 May 24.

31 Curtis, L, Burns A 2017 Unit Costs of Health and Social Care 2017, Personal Social Services Research Unit, University of Kent, Canterbury. Available from: http://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2017/

32 Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B Valuing health-related quality of life: An EQ-5D-5L value set for England; Health Economics. 2018;27:7–22; DOI: 10.1002/hec.3564

33 Hunter RM, Baio G, Butt T, Morris S, Round J, Freemantle N An Educational Review of the Statistical Issues in Analysing Utility Data for Cost-Utility Analysis. PharmacoEconomics 2015;33 (4):355-366.