Primary Care Clinical Trials Unit

STATISTICAL ANALYSIS REPORT

OxLith

OxLith: Exploration of the short-term physical and psychological effects of lithium in mood instability

	NAME	TITLE	Signature	Date
Written by:	Alecia Nickless	Senior Trial Statistician		
Reviewed by:		Senior Trial statistician		
Approved by:	John Geddes	Chief Investigator		

Version History

Version:	**Version Date:**	**Changes:**
0.1	9 November 2018	Original version
0.2	9 November 2018	Excluded 2 participants from the primary analysis who did not report compliance at 6 weeks.

TABLE OF CONTENTS

LIST OF TABLES

1 INTRODUCTION

The main focus of OxLIth is the effect of lithium on mood variability measured using validated scales of affective symptoms.

In addition to causing mood variability, bipolar disorder is known to affect cognitions, neural pathways, behaviour (including activity levels and sleep patterns) and circadian rhythms. Exploration of the effects of lithium on these parameters and the correlation between these effects and changes in mood form a secondary focus for the trial. In addition to this, OxLith will look for changes to renal, thyroid and parathyroid function and to inflammatory markers that might be caused by treatment with lithium.

The investigations have been grouped into five main overlapping themes exploring the effects of lithium on: 1) mood instability; 2) cognition and neural pathways; 3) behaviour; 4) circadian rhythms and 5) physiological measures.

Trial/Study statistician(s):

Alecia Nickless

Validation statistician(s):

Name and contact details

Chief Investigator:

Prof John Geddes

Trial/Study Manager:

Kate Saunders

Data Manager:

Jennifer Rendell

1.1 VALIDATION

Validation of results presented in this report was conducted by {name of statistician}. All results/major endpoints/primary endpoint were validated by independent programming using {name of program}. Results from Stata output were checked for transcription errors. Further details of validation including validation programs are saved on the PC-CTU restricted drive in the project folder in the subfolder "STATS\4. Analysis\6.Validation -Name of Validater".

1.2 SOFTWARE EMPLOYED

STATA 15.2 was used to execute the analyses.

2 METHODS

2.1 BACKGROUND INFORMATION

Bipolar Disorder affects over 2% of the world's population [Merikangas, 2011]. The onset of symptoms typically appears in late adolescence or early adulthood and the risk of developing new episodes remains relatively high for at least 40 years [Angst, 2003]. Furthermore, the lifetime risk of suicide for people with bipolar disorder is over 10 times higher than in the general population and the reduction in life expectancy between 9 and 20 years [Chesney, 2014].

Historically bipolar disorder was viewed as an illness characterised by episodes of depression and (hypo)mania interspersed with periods of euthymic (stable) mood with the focus of treatment being the resolution and prevention of episodes. More recently observational studies [Judd, 2002; Judd, 2003] and the use of prospective monitoring of patient self-reports of mood [Bauer, 2004; Bopp, 2010] have shown the prevalence of chronic inter- as well as intra-episode mood instability which is highly associated with the long-term social and functional impairment in bipolar disorder. Furthermore, subthreshold instability is often present long before the emergence of a first clinical mood episode making it an appropriate target for the search for an understanding of basic disease mechanisms and effects of treatment.

In 2008, a panel of experts in the study of antidepressant drugs highlighted problems conducting clinical trials in major depressive disorder arising from weaknesses in both diagnosis and clinical outcome measures [Gelenberg, 2008]. In particular, they expressed concerns that illness was defined phenomenologically rather than biologically and that there was a lack of surrogate markers of disease and treatment response. The lack of knowledge of basic disease mechanisms and consequent absence of validated pharmacological targets has similarly prevented substantial progress in treatment discovery in bipolar disorder [Geddes and Miklowitz, 2013].

Most treatments for mania and depression and for prevention of further episodes were developed for other indications (including epilepsy, depression, and psychosis). As a result, much of the clinical evidence has been obtained from patients with other conditions and use of treatments in bipolar disorder based on apparent similarities between indications. Lithium remains the only therapy primarily developed for treatment of bipolar disorder. It has increasingly been used as comparator in recent trials and is now convincingly established as an effective drug with a protective effect against suicide in mood disorders – reducing the risk by more than 80% [Cipriani, 2013]. For a proportion of patients, lithium provides very effective and tolerable symptom control. However, lithium treatment has a narrow therapeutic range and is associated with a number of adverse effects including those affecting renal, thyroid and parathyroid function [McKnight, 2012]). Better characterisation of the mechanisms of action of lithium is essential to inform both its current use for treatment of bipolar disorder and the generation of new targets for rational drug development.

2.2 TRIAL/STUDY DESIGN

Randomised, 6-week, double-blind, placebo-controlled trial. During the trial, routine psychiatric care will be provided as required by psychiatrists from the Bipolar Disorder Research Clinic (BDRC) based at the NIHR-CRF. At the end of the trial, participants will return to routine care i.e. continued psychiatric care at the Research Clinic, referral to a Community Mental Health Team or return to Primary Care as appropriate.

2.2.1 PRE-RANDOMISATION PHASE

Participants who give Informed Consent and have a diagnosis of both bipolar disorder (assessed using the Structured Clinical Interview for Axis I Disorders SCID-I) and current mood instability will enter a pre-randomisation phase. They will be asked to complete a battery of cognitive tests, be provided with an iPad and activity monitor(s) and be set up to provide information about their mood, cognitions and activities throughout the trial, including registration on True Colours. Vital signs, height, weight, blood pressure, pulse and temperature will be measured, blood samples collected, an ECG performed and, for women of child bearing potential (regardless of current use of contraception), a urine pregnancy test will be done. The pre-randomisation phase will enable participants to become familiar with True Colours, PANAS, OxLeq ratings and completing the cognitive tests prior to randomisation. At this visit participants will be fitted with an ePatch and asked to wear it for 72-hours and to take cheek swabs and saliva samples 4-hourly over a subsequent 32-hour period. The duration of this phase will be approximately 2 weeks. If randomisation is delayed for any reason the participant may be withdrawn from the trial and, if appropriate, be re-screened at a future date when baseline measurements that are subject to change will be re-measured.

2.2.2 RANDOMISATION VISIT

At the end of the pre-randomisation phase, participants will attend a randomisation visit when they will be randomly allocated oral lithium or matched placebo tablets in the ratio of 1:1. Prior to randomisation the psychiatrist will complete the CGI-BP scales to rate severity of illness and the participant will be asked to complete the Sleep Condition Indicator (CSI) questionnaire. The randomised phase will last for 6-weeks (including titration phase). At this visit appointments will be made for an MRI scan and, if a scanner is available, an MEG scan during week 4.

2.2.3 RANDOMISED PHASE VISITS

During the randomised phase, participants will be asked to attend brief assessment visits at 4-days, 8-days and a longer visit in the fourth week post-randomisation. If required, additional visits will be arranged between these times. Lithium levels and reported adherence to treatment will be checked at each visit and reviewed by a psychiatrist who will recommend adjustment of doses where required. Additional supplies of lithium/placebo will be provided as needed and participants will be asked about any Adverse Events (AEs).

At the 8-day visit the psychiatrist will make a clinical assessment of severity of illness and of change in mood symptoms since the start of the randomised phase using the CGI-BP and participants will be provided with a second set of sample tubes for cheek swabs and cortisol and melatonin collection for use just before the 4-week visit.

Brain scans will take place in the fourth week after randomisation, a second battery of cognitive tests will be completed and the psychiatrist will complete the CGI-BP for the period since the 8-day visit. Participants will be fitted with an ePatch (Appendix H) for the second 72-hour measurement period.

Participants will also be asked to attend a final 6-week visit when vital signs will be checked, the ECG, SCI and CGI-BP repeated for change since randomisation and since the 4-week visit and blood samples collected. Analyses will include lithium level and thyroid, parathyroid, renal function and levels of inflammatory markers. Checks of adherence and for adverse events will be carried out at this visit. Also at this visit participants will be

informed of their treatment allocation and discuss with their psychiatrist from the BDRC options for ongoing treatment and clinical care (see 8.10).

2.2.4 POST-RANDOMISED PHASE FOLLOW-UP

Participants will be contacted following cessation of trial treatment to review the status of any reported AEs and to check for any unreported AEs.

2.3 OBJECTIVES

2.3.1 PRIMARY OBJECTIVE

1. To compare the effects of lithium and placebo on mood symptoms over 6 weeks.

2.3.2 SECONDARY OBJECTIVES

2. To compare the effects of lithium and placebo on measures of cognitive function and to correlate these with current mood.
3. To compare the effects of lithium and placebo on variability in neural dynamics during MRI and, when available, MEG scans. This will include scans during resting state and whilst performing neuropsychological tasks.
4. To explore the effects of lithium on physical activity, sleep and (optional) social interactions.
5. To develop a profile of the circadian system in people with bipolar disorder with and without lithium treatment.
6. To explore of short-term effects of lithium on:
    a. thyroid function
    b. parathyroid function
    c. renal function
    d. Inflammatory markers
    e. Heart rate variability

2.4 TARGET POPULATION

INCLUSION CRITERIA

- Willing and able to give informed consent to participate in the trial
- Male or female
- Aged 18 or over
- Meeting criteria for bipolar disorder
- Clinical complaint of significant mood instability
- Clinical uncertainty about the prescription of lithium
- No clear indication for alternative treatment
- Pre-treatment tests including renal, cardiac, thyroid and parathyroid functions acceptable for initiation of treatment with lithium
- Willing and able to comply with all trial requirements including mood and behavioural monitoring (True Colours) and MRI and MEG scanning and blood tests (assessed by a psychiatrist).

- Willing to allow his/ her General Practitioner and consultant, if appropriate, to be notified of his/her participation in the trial.

3.2.2 Exclusion criteria

- Contraindication(s) to lithium (as documented in the Summary of Product Characteristics for Priadel)
- Currently taking any psychotropic drug that cannot be withdrawn (i.e. antidepressant, antipsychotic, mood stabiliser, benzodiazepine, non-benzodiazepine sleeping tablets) including as required (prn) medication**
- Clinically significant alcohol or substance use
- Requiring immediate treatment for an acute mood episode such that placebo would be inappropriate
- Clinical indication for immediate treatment with lithium
- Female and pregnant, lactating or planning a pregnancy during the course of the trial
- Female of child-bearing potential not willing to use effective contraception
- Participation in another research trial involving an investigational medicinal product in the past 12 weeks
- Judged to be at significant immediate risk of suicide/self-harm

Plus

- Patients who have a pacemaker, non-MR-compatible metal implant, or any other contraindication for MR or MEG brain scanning will be excluded from the corresponding brain scanning element(s) of Theme 2.

* Patients with a primary diagnosis of bipolar disorder with co-morbid anxiety or borderline personality disorder are not excluded.

** No patient will be withdrawn from effective medication or treatment for the purposes of this trial

2.5 Interventions

Each participant will be randomised to either lithium or placebo with participants, clinicians and researchers involved in trial recruitment and assessment visits being blind to allocation. Unless there are clinical reasons to start at a lower dose (which could include low body weight and previous experience of adverse effects of medications), participants will be prescribed lithium/placebo at an initial dose of 400mg/day to be taken at night.

2.6 Outcomes measures

2.6.1 Primary outcome

The primary outcome measures will be a change from baseline in the metric at 6 weeks from the date of randomisation. For assessment of QIDS-SR16, ALTMAN and CGI-BP version scores which are closest to 6 weeks from randomisation and within the range of 5-7 weeks inclusive, shall be used in the analysis. If two assessments are equidistant from 6 weeks then the later measurement will be considered the primary endpoint. Follow-up measurements are also available at 8 days post randomisation and 4 weeks post randomisation.

PANAS ratings were assessed daily. The PANAS positive and negative affect scores will be extracted closest to baseline, eight days follow-up, 4 weeks and 6 weeks follow-up.

1. Change in QIDS-SR_{16} score between baseline and 6 weeks.

   QIDS-SR_{16} is the Quick Inventory of Depressive Symptomatology. It is made up of 16 items.

   Questions correlate with the nine symptom criterion domains, which include: sleep disturbance (initial, middle, and late insomnia or hypersomnia) (Questions 1-4), sad mood (Question 5), decrease/increase in appetite/weight (Questions 6-9), concentration (Question 10), self-criticism (Question 11), suicidal ideation (Question 12), interest (Question 13), energy/fatigue (Question 14), psychomotor agitation/retardation (Questions 15-16). The following 10 scores are derived. 1.) the highest score of any 1 of the 4 sleep items (1-4), 2.) score of item 5, 3.) the highest score of any 1 of the appetite/weight items (6-9), 4.) score of item 10, 5.) score of item 11, 6.) score of item 12, 7.) score of item 13, 8.) score of item 14, 9.) highest score on either of the psychomotor items (items 15-16), and 10.) Sum all nine of these items to give the total score which can range between 0 and 27.

   Severity of depression is based on the total score as 1.) No depression (1-5), 2.) mild depression (6-10), 3.) moderate depression (11-15), 4.) severe depression (16-20), 5.) very severe depression (21-27).

2. Change in ALTMAN score between baseline and 6 weeks.

   ALTMAN mania scale - 5 questions, scored from 0 to 4. Scores are summed to obtain overall score which can range from 0 to 20, with higher values indicating increasing probability of a manic or hypomanic condition.

3. Change in PANAS (positive affect and negative affect) ratings between baseline and 6 weeks.

   The Positive and Negative Affect Schedule (PANAS) comprises two mood scales, one that measures positive affect and the other which measures negative affect. These measures are treated separately.. Participants in the PANAS are required to respond to a 20-item test using 5-point scale that ranges from very slightly or not at all (1) to extremely (5). The words are: 1.) interested, 2.) distressed, 3.) excited, 4.) upset, 5.) strong, 6.) guilty, 7.) scared, 8.) hostile, 9.) enthusiastic, 10.) proud, 11.) irritable, 12.) alert, 13.) ashamed, 14.) inspired, 15.) nervous, 16.) determined, 17.) attentive, 18.) jittery, 19.) active, 20.) afraid. To get the positive affect score add the scores on items 1.) interested, 3.) excited, 5.) strong, 9.) enthusiastic, 9.) enthusiastic, 12.) alert, 14.) inspired, 16.) determined, 17.) attentive, 19.) active. To get the negative affect score add the scores on items 2.) distressed, 4.) upset, 6.) guilty, 7.) scared, 8.) hostile, 11.) irritable, 13.) ashamed, 15.) nervous , 18.) jittery, 20.) afraid. Positive and negative affect scores can range between 10 and 50.

4. Change in CGI-BP version between baseline and 6 weeks.

   The Clinical Global Impressions (CGI)-BP is an observer-rated scale that measures illness severity. The CGI is rated on a 7-point scale, with the severity of illness scale using a range of responses from 1 (normal) through to 7 (amongst the most severely ill patients). Each component of the CGI is rated separately; the instrument does not yield a global score. CGI Severity Scale is a 7-point scale that requires the clinician to rate the severity of the patient's illness at the time of assessment, relative to the clinician's past experience with patients who have the same diagnosis. Possible ratings are: 1.) normal, not at all ill 2.) borderline mentally ill 3.) mildly ill 4.) moderately ill 5.) markedly ill 6.) severely ill 7.) among the most extremely ill patients. The

occurrence of manic and depressive episodes and their severity were assessed using the CGI-BP, considering the highest severity score as the final result. If no episode is recorded the CGI-BP severity scale will be set at 1.

2.6.2 SECONDARY OUTCOMES

1. Change in the cognitive assessment PANAS rating over time.
2. Measures of variability in performance across time in the PANAS rating.
3. Blood oxygen level dependent signal during rest and task performance; induced and evoked field activity. *The analysis of this outcome will be reported in a separate report.*
4. Frequency and amplitude of movements during daytime activity. *The analysis of this outcome will be reported in a separate report.*
5. Categorisation of activity types. *The analysis of this outcome will be reported in a separate report.*
6. Duration, timing, and quality of sleep. *The analysis of this outcome will be reported in a separate report.*
7. Change in sleep pattern (SCI) pre- and during randomised phase.
8. Changes in gene expression, cortisol and melatonin levels. *The analysis of this outcome will be reported in a separate report.*
9. Changes in blood levels of relevant biomarkers including:
    a) T4 (thyroxine), T3 (triiodothyronine), TSH (a thyroid stimulating hormone) and thyroid antibodies.
    b) Calcium, PTH (parathyroid hormone), and vitamin D.
    c) Creatinine, cystatin C and NGAL (neutrophil gelatinase-associated lipocalin).
    d) C-reactive protein (CRP) and interleukin 6 (IL-6).
10. Change in heart rate variability over 72 hours (using ePatch). *The analysis of this outcome will be reported in a separate report.*

2.7 SAMPLE SIZE

The target sample size is 40 randomised participants (20 of whom will be allocated treatment with lithium and 20 with placebo).

In an earlier study (OXTEXT-6 - http://oxtext.psych.ox.ac.uk/) the mean ND for the QIDS-SR16 was 1.98 (standard deviation 1.73). Considering a 25% reduction in ND as a minimum clinically meaningful change over 6 weeks, two pre-randomisation measurements (correlation 0.4) a total sample size of 40 randomised participants would provide 81% power. Two post randomisation measurements would provide 94% power (86% if correlation is 0.2; 81% if 0.1). Similar power will be achieved for the ALTMAN Mania scale.

2.8 RANDOMISATION AND BLINDING IN THE ANALYSIS STAGE

2.8.1 RANDOMISATION VISIT

The randomisation visit will take place approximately 2 weeks after trial entry.

During this visit eligibility for the randomised phase will be checked including continued consent, absence of contraindications to lithium and adherence to mood ratings, cognitive tasks, the carrying of the activity monitor(s), return of the ePatch and the provision of cheek swabs and saliva samples. Participants will be

asked to complete the sleep questionnaire (CSI). Any difficulties will be explored and, if possible, overcome to enable the participant to continue in the trial.

Eligible participants will be entered into the randomised phase (see below) and provided with trial medication and date(s) for future visits including for MRI and, when available MEG, scans.

2.8.2 Randomisation and allocation concealment

Randomisation will be done in real time centrally via the internet according to an algorithm that minimises on two prognostic factors. Each participant will be randomised to either lithium or placebo with participants, clinicians and researchers involved in trial recruitment and assessment visits being blind to allocation.

The computer-generated randomisation schedule will be designed by the software development team based at the University of Oxford Department of Psychiatry. Central randomisation will ensure full allocation concealment [Schulz 1995]. A non-deterministic minimisation algorithm will be used to produce treatment groups balanced for important prognostic factors. The first 10 participants will be allocated treatment randomly without minimisation to avoid predictability. Subsequently, the minimisation algorithm will be applied with an allocation ratio that is not fully deterministic: there will be an 80% bias in favour of allocations that minimise the imbalance.

2.8.3 Minimisation variables

The randomisation algorithm will minimise separately on two variables: age (<25 and ≥ 25) and gender (M, F) which are related to prognosis.

2.8.4 Unblinding of treatment

A list of current participants showing their allocation will be maintained by nurses/research assistants from the NIHR-CRF who are designated unblind.

Emergency unblinding: Participants will be given a card to carry to say that they are in the trial and may be taking lithium or matched placebo. Lithium is widely used and therefore any medical doctor attending a trial participant who is experiencing a medical emergency will be fully aware of its clinical effects. In the event of a medical emergency, routine care for a patient who may have been taking lithium and has symptoms consistent with lithium toxicity would be to obtain a lithium level immediately thus negating the need for the treating doctor to be informed about the trial allocation. No arrangements for out of hours unblinding will therefore be made.

Non-emergency unblinding: Allocation of treatments will be recorded on a Randomisation List which will be updated when each new participant enters the randomised phase. The list will be held by unblind nurses/research assistants from the NIHR-CRF whose only involvement with the trial will be holding the list and managing the provision of sham blood results (see 8.9). These staff will, on the instructions of the Chief Investigator or delegate, access the list and reveal the allocation to an individual participant and to the participant's psychiatrist.

The analysis was performed unblinded to treatment allocation. The statistical analysis plan was completed and signed off prior to data lock for final analysis.

2.9 Analysis for Data Monitoring Committee meetings

No interim analysis has been performed on the data.

2.10 Definition of population for analysis

The primary statistical analyses were per protocol with data for all participants being included up to the point where they stop trial treatment or complete the randomised phase. The intention to treat analysis (ITT) population included all patients randomised irrespective of whether they received any trial medication, and who provided outcome data as required for any of the endpoints listed. The ITT analysis would have been performed as a sensitivity analysis of the primary analyses, but all participants who provided outcome data reported 65% or more adherence to the trial medication regime, and therefore the per protocol and ITT populations were the same.

2.11 Deviation from SAP

State any analyses deviated from SAP or additional analyses not specified in the SAP

3 Results

3.1 Representativeness of Study Sample and Patient Throughput

A CONSORT flow diagram of the participants through the study is provided in Appendix I. A total of 40 participants were screened, and 35 participants proceeded to randomisation of treatment with Lithium or placebo. A total of 19 participants were randomised to the Lithium arm and 16 to the placebo arm. One participant was withdrawn from the study before the Day 8 follow-up due to non-compliance.

3.2 Recruitment

The first participants entered the study on the 15th August 2015. The last participant was randomised on the 15th of January 2018.

3.3 BASELINE CHARACTERISTICS OF PARTICIPANTS

TABLE 1 BASELINE CHARACTERISTICS

	N	LITHIUM	N	PLACEBO	N	OVERALL
DEMOGRAPHICS						
Age, mean(SD) [range]	19	28.9 (9.8) [18.5; 63.1]	16	35.2 (13.8) [19.2; 62.6]	35	31.8 (12.0) (18.5; 63.1)
Gender *Male, n(%)* *Female, n(%)*	19	 8 (42.11%) 11 (57.9%)	16	 7 (56.3%) 9 (43.8%)	35	 15 (42.9%) 20 (57.1%)
Ethnic group *White British, n(%)* *White Other, n(%)* *Mixed, n(%)* *Asian, n(%)* *Black American, n(%)* *Other, n(%)*	19	 13 (68.4%) 1 (5.3%) 1 (5.3%) 0 (0.0%) 1 (5.3%) 3 (15.8%)	16	 9 (56.3%) 1 (6.3%) 1 (6.3%) 1 (6.3%) 0 (0.0%) 4 (25.0%)	35	 22 (62.9%) 2 (5.7%) 2 (5.7%) 1 (2.9%) 1 (2.9%) 7 (20.0%)
ASSESSMENTS						
Bipolar diagnosis *BD I* *BD II* *BD*	19	 3 (15.8%) 16 (84.2%) 0 (0.0%)	16	 4 (25.0%) 11 (68.8%) 1 (6.3%)	35	 7 (20.0%) 27 (77.1%) 1 (2.9%)
QIDS (Quick Inventory of Depressive Symptomatology), mean(SD) [range]	19	10.6 (6.6) [1; 25]	16	11.6 (5.4) [3; 21]	35	11.1 (6.0) [1; 25]
Altman (Altman Self Rating Scale for Mania), mean (SD) [range]	19	2.3 (3.7) [0, 12]	16	3.6 (3.9) [0; 14]	35	2.9 (3.8) [0; 14]
PANAS + (Positive and Negative Affect Scale Positive subscale), mean (SD) [range]	19	6.8 (4.7) [0; 14]	16	7.1 (5.2) [1; 16]	35	6.9 (4.8) [0; 16]
PANAS - (Positive and Negative Affect Scale Negative subscale), mean (SD) [range]	19	5.0 (4.5) [0; 13]	16	5.5 (5.9) [0; 17]	35	5.2 (5.1) [0; 17]
CGI (Clinical Global Impression) Mania, median (IQR) [range]	15	1 (1; 2) [1; 2]	11	1 (1; 2) [1; 4]	26	1 (1; 2) [1; 4]
CGI Depression, Median (IQR) [range]	15	2 (1; 4) [1; 4]	11	3 (2; 3) [1; 4]	26	2 (1; 3) [1; 4]
CGI Bipolar, median (IQR) [range]	15	2 (2; 4) [1; 4]	11	2 (2; 3) [1; 4]	26	2 (2; 3) [1; 4]

	N	LITHIUM	N	PLACEBO	N	OVERALL
BLOOD TESTS mean(SD) [range]						
IL6	18	1.4 (0.6) [0.66; 3.06]	14	1.4 (0.6) [0.77; 2.80]	32	1.4 (0.6) [0.66; 3.06]
NGAL	18	0.13 (0.03) [0.09; 0.21]	14	0.13 (0.03) [0.07; 0.19]	32	0.13 (0.03) [0.07; 0.21]
T4 Thyroxine	19	12.9 (1.5) [11.0; 16.9]	14	13.1 (0.8) [11.4; 14.5]	33	13.0 (1.3) [11.0; 16.9]
T3 Triiodothyronine	19	4.6 (0.6) [3.5; 5.6]	15	4.3 (0.4) [3.6; 5.1]	34	4.5 (0.5) [3.5; 5.6]
TSH Thyroid stimulating hormone	19	1.3 (0.6) [0.31; 2.83]	16	1.3 (0.7) [0.17; 3.05]	35	1.3 (0.6) [0.17; 3.05]
Calcium	19	2.4 (0.1) [2.17; 2.53]	16	2.4 (0.1) [2.23; 2.60]	35	2.4 (0.1) [2.17; 2.60]
PTH parathyroid hormone	17	5.3 (1.5) [3.2; 8.0]	16	5.4 (2.9) [2.4; 12.9]	33	5.4 (2.2) [2.4; 12.9]
Vitamin D	17	45.9 (18.7) [12; 89]	14	53.7 (35.7) [11; 147]	29	49.6 (27.8) [11; 147]
Creatinine	19	68.1 (14.1) [51; 98]	16	67.3 (13.1) [50; 92]	35	67.7 (13.5) [50; 98]
Cycstatin C	17	0.90 (0.13) [0.65; 1.06]	16	0.87 (0.12) [0.68; 1.19]	33	0.89 (0.12) [0.65; 1.19]
CRP C-reactive protein	16	1.9 (3.1) [0.3; 13.0]	14	2.5 (3.4) [0.3; 11.4]	30	2.2 (3.2) [0.3; 13.0]

3.4 NUMBER ANALYSED

One participant in the placebo arm was withdrawn shortly after baseline due to non-compliance. All remaining participants were followed up to the primary endpoint (6 weeks), but not all outcomes were available for each participant. The availability of the primary outcomes is provided in the CONSORT diagram in Appendix 1.

3.5 Primary Analyses

3.5.1 Primary outcomes

TABLE 2 SUMMARY STATISTICS AND ADJUSTED TREATMENT DIFFERENCE IN THE QIDS, ALTMAN, AND PANAS RATING SCALES. THE ANALYSIS WAS PERFORMED ON THE PER PROTOCOL POPULATION AND THE PRIMARY ENDPOINT IS 6 WEEKS.

Outcome	N	Lithium		N	Placebo		Treatment Difference* (95% CI); Standardised Difference	p-value
		Mean (sd)	Mean Change (sd)		Mean (sd)	Mean Change (sd)		
QIDS Total								
Baseline	18	11.0 (6.6)		14	11.4 (5.5)			
8 Days	18	9.9 (5.8)	-1.06 (4.39)	14	11.6 (5.3)	0.29 (4.98)	-1.16 (-4.48; 2.17)	0.4960
4 Weeks	18	7.6 (4.4)	-3.44 (5.61)	14	11.1 (6.6)	-0.21 (5.18)	-3.04 (-6.37; 0.28)	0.0729
6 Weeks	18	9.1 (6.0)	1.94 (7.01)	13	8.9 (6.1)	-1.92 (5.81)	0.45 (-2.96; 3.86)	0.7978
ALTMAN Total								
Baseline	18	1.8 (2.9)		14	3.4 (4.1)			
8 Days	18	2.6 (4.2)	0.78 (3.42)	14	2.2 (2.8)	-1.14 (5.48)	0.74 (-1.63; 3.11)	0.5425
4 Weeks	18	3.6 (3.8)	1.78 (4.71)	14	2.7 (4.3)	-0.64 (2.68)	1.24 (-1.13; 3.61)	0.3066
6 Weeks	18	2.7 (3.0)	0.94 (3.57)	13	3.2 (3.4)	-0.46 (3.28)	0.01 (-2.44; 2.46)	0.9911
PANAS +								
Baseline	18	6.6 (4.7)		14	7.9 (4.9)			
8 Days	18	6.4 (4.8)	-0.11 (5.38)	14	5.1 (5.0)	-2.79 (6.89)	1.03 (-1.92; 3.98)	0.4954
4 Weeks	16	5.3 (5.3)	-1.13 (5.88)	14	6.9 (5.3)	-1.0 (5.01)	-1.84 (-4.86; 1.17)	0.2310
6 Weeks	16	4.0 (3.8)	-2.44 (4.83)	12	8.0 (5.1)	0.5 (3.66)	-4.10 (-7.16; -1.04)	0.0086
PANAS –								
Baseline	18	5.0 (4.7)		14	5.1 (5.2)			
8 Days	18	4.0 (4.5)	-1.00 (4.92)	14	6.5 (5.5)	1.43 (5.51)	-1.35 (-4.45; 1.76)	0.3950
4 Weeks	16	2.6 (4.3)	-2.69 (5.19)	14	5.1 (5.2)	0.07 (5.50)	-1.25 (-4.39; 1.90)	0.4368
6 Weeks	16	4.1 (5.2)	-1.19 (5.90)	12	3.6 (4.6)	-1.75 (4.61)	1.24 (-1.94; 4.41)	0.4449

***Linear mixed effects model with fixed effects for baseline total score, gender, age, assessment point (8 days, 4 weeks or 6 weeks), randomised group (Lithium or placebo), and the interaction between assessment point and randomised group, and a random effect for participant.**

TABLE 3 SUMMARY STATISTICS AND GROUP COMPARISONS IN THE CGI ASSESSMENT. THE ANALYSIS WAS PERFORMED ON THE PER PROTOCOL POPULATION AND THE PRIMARY ENDPOINT IS 6 WEEKS.

Outcome	N	Lithium		N	Placebo		p-value* for difference in change in CGI severity score
		Median[£] (IQR)	Median Change[$] (IQR)		Median[£] (IQR)	Median Change[$] (IQR)	
CGI Mania							
Baseline	14	1 (1; 2)		10	1 (1; 2)		
8 Days	14	1 (1; 1)		10	1 (1; 1)		
4 Weeks	14	1 (1; 1)	4 (4; 4)	10	1 (1; 1)	4 (4; 4)	0.9690
6 Weeks	14	1 (1; 1)	4 (4; 4)	10	1 (1; 2)	4 (4; 4)	0.4453
6 – 4 Weeks	13		4 (3; 4)	10		4 (4; 5)	0.2165
CGI Depression							
Baseline	14	2 (1; 4)		10	2.5 (2; 3)		
8 Days	14	2 (1; 2)		10	2 (1; 3)		
4 Weeks	14	1 (1; 3)	4 (3; 4)	10	1.5 (1; 3)	3.5 (3; 4)	0.4554
6 Weeks	14	2 (1; 3)	3.5 (2; 4)	10	1 (1; 2)	3 (2; 4)	0.9047
6 – 4 Weeks	14		4 (2; 5)	10		3.5 (2; 4)	0.3346
CGI Bipolar							
Baseline	14	2 (2; 4)		10	2 (2; 3)		
8 Days	14	2 (1; 3)		10	2 (1; 3)		
4 Weeks	14	2 (1; 3)	4 (4; 5)	10	2 (1; 3)	3.5 (3; 4)	0.1816
6 Weeks	14	2 (1; 3)	3 (2; 4)	10	2 (1; 3)	3 (2; 4)	0.7410
6 – 4 Weeks	14		4 (2; 4)	10		4 (2; 4)	0.7320

***Wilcoxon signed rank test performed to test differences between placebo and Lithium groups.**

[£] Response to "Considering your total clinical experience with bipolar patients, how severely ill is the patient at this time"

(1 = Normal, not ill; 2 = Minimally ill; 3 = Mildly ill; 4 = Moderately ill; Markedly ill)

[$] Response to "Compared to the trial visit at Randomisation (previous visit), how much has the patient changed?"

(1 = Very much improved; 2 = Much improved; 3 = Minimally improved; 4 = No change; 5 = Minimally worse; 6 = Much worse)

3.6 SECONDARY ANALYSES

3.6.1 SECONDARY OUTCOMES

TABLE 4 SUMMARY STATISTICS AND ADJUSTED TREATMENT DIFFERENCES FOR THE PANAS OUTCOME

Outcome	Lithium (N = 19)	Placebo (N = 15)	Treatment Difference* (95% CI); Standardised Difference	p-value
	Mean (95% CI)	Mean (95% CI)		
PANAS +				
Baseline	5.7 (4.3; 7.2)	4.7 (3.1; 6.3)		
8 Days	5.5 (4.1; 6.9)	5.3 (3.8; 6.9)	0.17 (-1.95; 2.30)	0.8734
4 Weeks	5.0 (3.6; 6.4)	6.9 (5.4; 8.5)	-1.93 (-4.04; 0.18)	0.0734
6 Weeks	4.7 (3.2; 6.1)	8.1 (6.4; 9.7)	-3.40 (-5.63; -1.17)	0.0028
PANAS -				
Baseline	3.2 (1.7; 4.8)	5.8 (4.1; 7.5)		
8 Days	3.3 (1.8; 4.8)	5.4 (3.7; 7.0)	-2.05 (-4.34; 0.24)	0.0795
4 Weeks	3.6 (2.1; 5.1)	4.3 (2.6; 5.9)	-0.70 (-2.97; 1.58)	0.5487
6 Weeks	3.7 (2.2; 5.3)	3.5 (1.7; 5.2)	0.25 (-2.12; 2.62)	0.8363

***Linear mixed effects model with fixed effects for baseline total score, gender, age, continuous time, randomised group (Lithium or placebo), pre or post randomisation period, and the interaction between time and randomised group, between time and pre or post randomisation period and between pre or post randomisation period and randomised group, and the three way interaction between pre or post randomisation period, time and randomised group, and with a random effect for participant.**

TABLE 5 SUMMARY STATISTICS AND ADJUSTED TREATMENT DIFFERENCES FOR THE STABILITY MEASURES FOR THE PANAS POSITIVE AND NEGATIVE AFFECT SCORES

Outcome	N	Lithium		N	Placebo		Treatment Difference* (95% CI); Standardised Difference	p-value
		Mean (sd)	Mean Change (sd)		Mean (sd)	Mean Change (sd)		
PANAS + standard deviation (sd)	17			13				
Baseline		2.7 (1.8)			2.8 (1.2)			
6 Weeks		1.9 (1.2)	-0.73 (1.67)		2.1 (1.5)	-0.67 (1.87)	-0.15 (-1.21; 0.90)	0.7689
PANAS - standard deviation (sd)	17			13				
Baseline		2.7 (1.4)			2.9 (1.4)			
6 Weeks		2.1 (1.5)	-0.56 (1.68)		1.6 (1.3)	-1.31 (1.55)	0.62 (-0.49; 1.74)	0.2609
PANAS + root mean square of successive differences (rmssd)	17			13				
Baseline		3.2 (2.0)			3.5 (1.5)			
6 Weeks		2.6 (1.7)	-0.53 (2.08)		2.5 (1.9)	-0.94 (2.06)	0.19 (-1.18; 1.56)	0.7824
PANAS - root mean square of successive differences (rmssd)	17			13				
Baseline		3.3 (2.2)			3.5 (1.5)			
6 Weeks		2.6 (1.8)	-0.75 (2.51)		1.94 (1.8)	-1.59 (1.81)	0.60 (-0.75; 1.96)	0.3685
PANAS + entropy	14			12				
Baseline		2.0 (0.5)			2.2 (0.4)			
6 Weeks		1.6 (0.9)	-0.33 (0.69)		2.2 (0.4)	0.05 (0.65)	-0.53 (-1.08; 0.03)	0.0619
PANAS - entropy	16			10				
Baseline		1.6 (0.7)			2.1 (0.4)			
6 Weeks		1.3 (0.8)	-0.22 (0.92)		1.9 (0.8)	-0.22 (0.83)	-0.45 (-1.27; 0.37)	0.2660

***Linear regression model with covariates for baseline measurement, gender, age, and randomised group (Lithium or placebo).**

TABLE 6 SUMMARY STATISTICS AND ADJUSTED TREATMENT DIFFERENCES FOR THE SECONDARY OUTCOMES

Outcome	N	Lithium		N	Placebo		Treatment Difference* (95% CI); Standardised Difference	p-value
		Mean (sd)	Mean Change (sd)		Mean (sd)	Mean Change (sd)		
SCI	14			10				
Baseline		16.4 (4.8)			12.4 (4.5)			
6 Weeks		16.3 (5.7)	-0.07 (5.21)		15.1 (7.1)	2.70 (7.59)	-0.28 (-6.29; 5.72)	0.9262
Blood tests								
IL6	18			14				
Baseline		1.4 (0.6)			1.4 (0.6)			
6 Weeks		1.7 (1.2)	0.34 (1.18)		1.4 (0.6)	0.01 (0.38)	-0.28 (-0.28; 0.83)	0.3271
NGAL	18			14				
Baseline		0.13 (0.03)			0.13 (0.03)			
6 Weeks		0.15 (0.03)	0.02 (0.02)		0.13 (0.04)	0.00 (0.03)	0.02 (0.00; 0.05)	0.0273
T4 Thyroxine	17			10				
Baseline		13.0 (1.6)			13.2 (1.0)			
6 Weeks		12.2 (1.8)	-0.81 (1.38)		12.8 (0.8)	-0.41 (0.76)	-0.32 (-1.32; 0.68)	0.5310
T3 Triiodothyronine	18			11				
Baseline		4.6 (0.6)			4.4 (0.4)			
6 Weeks		4.4 (0.7)	-0.14 (0.78)		4.5 (0.5)	0.10 (0.49)	-0.12 (-0.58; 0.33)	0.6018
TSH Thyroid stimulating hormone	18			12				
Baseline		1.2 (0.4)			1.4 (0.6)			
6 Weeks		2.2 (1.2)	1.03 (1.04)		1.6 (0.6)	0.13 (0.56)	0.63 (0.02; 1.24)	0.0418
Calcium	17			13				
Baseline		2.4 (0.1)			2.4 (0.1)			
6 Weeks		2.4 (0.1)	-0.01 (0.08)		2.4 (0.1)	0.00 (0.09)	-0.01 (-0.06; 0.05)	0.8057

Outcome	N	Lithium		N	Placebo		Treatment Difference* (95% CI); Standardised Difference	p-value
PTH parathyroid hormone	16			13				
Baseline		5.1 (1.4)			5.5 (3.1)			
6 Weeks		4.9 (1.4)	-0.21 (1.71)		5.0 (2.2)	-0.57 (1.46)	0.49 (-0.50; 1.49)	0.3243
Vitamin D	15			13				
Baseline		47.3 (19.4)			52.9 (36.6)			
6 Weeks		49.0 (21.4)	1.67 (22.83)		52.8 (31.1)	-0.15 (18.74)	7.4 (-5.3; 20.0)	0.2542
Creatinine	17			13				
Baseline		66.8 (14.3)			66.8 (13.5)			
6 Week		73.0 (14.4)	6.24 (4.78)		69.2 (12.8)	2.38 (6.41)	4.6 (-0.05; 9.2)	0.0525
Cycstatin C	14			12				
Baseline		0.90 (0.13)			0.89 (0.12)			
6 Weeks		0.94 (0.12)	0.05 (0.11)		0.90 (0.12)	0.00 (0.05)	0.06 (-0.02; 0.14)	0.1344
CRP C-reactive protein	13			10				
Baseline		1.2 (1.1)			2.4 (3.4)			
6 Weeks		5.2 (10.2)	4.01 (10.5)		2.7 (3.7)	0.28 (5.02)	1.0 (-6.5; 8.4)	0.7928

***Linear regression model with covariates for baseline measurement, gender, age, and randomised group (Lithium or placebo). Confidence intervals and p-values were obtained by bootstrapping due to skewness in the data.**

TABLE 7 COMPLIANCE TO STUDY MEDIATION REGIME REPORTED AT 4 DAYS AND AT 4 WEEKS

	Lithium (N=19)	Placebo (N=16)
Reported at 4 Days		
More than 65%	19 (100%)	15 (93.8%)
Between 50 and 65%	0 (0%)	1 (6.3%)
Reported at 4 Weeks		
More than 65%	19 (100%)	14 (87.5%)
Between 50 and 65%	0 (0%)	0 (0%)
Missing	0 (0%)	2 (12.5%)
Reported at 6 Weeks		
More than 65%	18 (94.7%)	14 (87.5%)
Between 50 and 65%	0 (0%)	0 (0%)
Missing	1 (5.3%)	2 (12.5%)

3.7 SENSITIVITY ANALYSES

TABLE 8 SUMMARY STATISTICS AND ADJUSTED TREATMENT DIFFERENCE IN THE QIDS, ALTMAN, AND PANAS RATING SCALES. THIS IS PERFORMED ON THE ITT POPULATION AND THE PRIMARY ENDPOINT IS 6 WEEKS.

Outcome	N	Lithium		N	Placebo		Treatment Difference* (95% CI); Standardised Difference	p-value
		Mean (sd)	Mean Change (sd)		Mean (sd)	Mean Change (sd)		
QIDS Total								
Baseline	19	10.6 (6.6)		15	11.1 (5.3)			
8 Days	19	9.7 (5.7)	-0.9 (4.3)	15	11.5 (5.2)	0.4 (4.8)	-1.4 (-4.4; 1.7);	0.3841
4 Weeks	19	7.5 (4.3)	-3.2 (5.6)	15	11.1 (6.4)	-0.07 (5.0)	-3.1 (-6.2; -0.04);	0.0473
6 Weeks	19	8.9 (5.8)	-1.7 (6.9)	14	8.7 (6.0)	-1.9 (5.6)	0.49 (-2.64; 3.63);	0.7577
ALTMAN Total								
Baseline	19	2.3 (3.7)		15	3.5 (4.0)			
8 Days	19	2.8 (4.2)	0.5 (3.5)	15	2.1 (2.8)	-1.4 (5.4)	0.9 (-1.4; 3.2);	0.4389
4 Weeks	19	3.4 (3.8)	1.1 (5.6)	15	2.5 (4.2)	-0.9 (2.8)	1.0 (-1.3; 3.2);	0.4091
6 Weeks	19	2.6 (3.0)	0.3 (4.6)	14	2.9 (3.4)	-0.8 (3.4)	-0.2 (-2.6; 2.1)	0.8594
PANAS +								
Baseline	19	6.8 (4.7)		15	7.5 (5.1)			
8 Days	19	6.6 (4.7)	-0.2 (5.2)	15	4.8 (5.0)	-2.7 (6.7)	1.7 (-1.1; 4.4);	0.2470
4 Weeks	17	5.5 (5.2)	-1.2 (5.7)	15	6.5 (5.4)	-1.0 (4.8)	-1.1 (-3.9; 1.8);	0.4542
6 Weeks	16	4.0 (3.8)	-2.4 (4.8)	12	8.0 (5.1)	0.5 (3.7)	-3.6 (-6.6; -0.6);	0.0177
PANAS –								
Baseline	19	5.0 (4.5)		15	4.7 (5.2)			
8 Days	19	4.1 (4.4)	-0.9 (4.8)	15	6.3 (5.4)	1.5 (5.3)	-1.5 (-4.4; 1.3)	0.2891
4 Weeks	17	2.8 (4.3)	-2.4 (5.1)	15	5.1 (5.0)	0.4 (5.4)	-1.5 (-4.4; 1.3)	0.2965
6 Weeks	16	4.1 (5.2)	-1.2 (5.9)	12	3.6 (4.6)	-1.8 (4.6)	1.1 (-1.9; 4.0)	0.4931

***Linear mixed effects model with fixed effects for baseline total score, gender, age, assessment point (8 days, 4 weeks or 6 weeks), randomised group (Lithium or placebo), and the interaction between assessment point and randomised group, and a random effect for participant.**

TABLE 9 SUMMARY STATISTICS AND GROUP COMPARISONS IN THE CGI ASSESSMENT. THE PRIMARY ENDPOINT IS 6 WEEKS.

Outcome	N	Lithium		N	Placebo		p-value* for difference in change in CGI severity score
		Median[£] (IQR)	Median Change[$] (IQR)		Median[£] (IQR)	Median Change[$] (IQR)	
CGI Mania							
Baseline	15	1 (1; 2)		10	1 (1; 2)		
8 Days	15	1 (1; 1)		10	1 (1; 1)		
4 Weeks	15	1 (1; 1)	4 (4; 4)	10	1 (1; 1)	4 (4; 4)	1.0000
6 Weeks	13	1 (1; 1)	4 (4; 4)	10	1 (1; 2)	4 (4; 4)	0.4453
6 – 4 Weeks	13		4 (3; 4)	10		4 (4; 5)	0.2165
CGI Depression							
Baseline	15	2 (1; 4)		10	2.5 (2; 3)		
8 Days	15	2 (1; 2)		10	2 (1; 3)		
4 Weeks	15	1 (1; 3)	4 (3; 4)	10	1.5 (1; 3)	3.5 (3; 4)	0.4233
6 Weeks	14	2 (1; 3)	3.5 (2; 4)	10	1 (1; 2)	3 (2; 4)	0.9047
6 – 4 Weeks	14		4 (2; 5)	10		3.5 (2; 4)	0.3346
CGI Bipolar							
Baseline	15	2 (2; 4)		10	2 (2; 3)		
8 Days	15	2 (1; 3)		10	2 (1; 3)		
4 Weeks	15	2 (1; 3)	4 (4; 5)	10	2 (1; 3)	3.5 (3; 4)	0.1739
6 Weeks	14	2 (1; 3)	3 (2; 4)	10	2 (1; 3)	3 (2; 4)	0.7410
6 – 4 Weeks	14		4 (2; 4)	10		4 (2; 4)	0.7320

***Wilcoxon signed rank test performed to test differences between placebo and Lithium groups.**

[£] Response to "Considering your total clinical experience with bipolar patients, how severely ill is the patient at this time"

(1 = Normal, not ill; 2 = Minimally ill; 3 = Mildly ill; 4 = Moderately ill; Markedly ill)

[$] Response to "Compared to the trial visit at Randomisation (previous visit), how much has the patient changed?"

(1 = Very much improved; 2 = Much improved; 3 = Minimally improved; 4 = No change; 5 = Minimally worse; 6 = Much worse)

3.8 SAFETY ANALYSES

No serious adverse events were reported.

TABLE 10 ADVERSE EVENTS REPORTED AT 4 DAYS AND AT 6 WEEKS

	Lithium (N=19)	Placebo (N=16)
Reported at 4 Days		
Yes	8 (42.1%)	6 (37.5%)
No	11 (57.9%)	10 (62.5%)
Reported at 6 Weeks		
Yes	10 (52.6%)	6 (35.5%)
No	8 (42.1%)	8 (50.0%)
Missing	1 (5.3%)	2 (12.5%)

3.9 ADDITIONAL EXPLORATORY ANALYSIS NOT SPECIFIED IN THE SAP

Any significant changes to the statistical analysis section in protocol or previous versions of SAP should be described here.

4 REFERENCES

Any relevant literature review referred to should be listed here.

5 APPENDICES

Appendix I. Flow diagram of trial participants