Statistical analysis plan

Author: Graham Dunn

Approved by the Data Monitoring and Ethics Committee (Chair: Stephen Cooper (Emeritus professor of Psychiatry, University of Belfast) and the Trial Steering Committee (Chair: John Geddes, University of Oxford).

The document below is as submitted except for minor formatting changes, without edits and with background taken from the protocol prior to the addition of new recruitment centres.

Data Analysis Plan

2nd draft, GD 13th May, 2016

EudraCT Number:	2010-022463-35
Sponsor:	Manchester Mental Health & Social Care Trust
Funder:	MRC/NIHR – EME Programme

Principal investigator

Bill Deakin

Trial co-ordinator

Emma Knox

Trial statistician

Graham Dunn

RESEARCH OBJECTIVES:

Research Questions

If minocycline is started early in schizophrenia does it improve negative symptoms or lessen their development over the next 12 months better than it does in established illnesses, and does this improve quality of life? Does minocycline work by its neuroprotective or anti-inflammatory actions or by its effects on glutamate?

Primary and subsidiary effectiveness predictions:

Hypothesis 1. Minocycline minimises later negative symptoms when administered during the acute phase of psychosis, compared to standard care alone.

Hypothesis 2. Minocycline reduces weight gain and adverse metabolic changes associated with standard antipsychotic treatments.

Hypothesis 3. Improvements in negative symptoms will translate into improved functioning and quality of life.

Mechanistic hypotheses:

Hypothesis 1. Minocycline works by lessening a degenerative process, which is most active in the acute phase of psychosis and is responsible for the development of negative symptoms. The hypothesis predicts that the loss of grey matter, known to occur during the early years following onset of psychosis, will be lessened by minocycline treatment and that this will correlate with and explain improved negative symptoms.

Hypothesis 2. Minocycline works by lessening an inflammatory process in the brain which gives rise to negative symptoms, possibly but not necessarily mediated by subtle neurodegeneration (see H1 above). The hypothesis predicts that circulating pro-inflammatory cytokines will be lessened by minocyline treatment.

Hypothesis 3. Minocycline works by ameliorating defective NMDA glutamate receptor function which mediates negative symptoms. The hypothesis predicts that minocycline will improve cortical function as measured by functional MRI (fMRI) activation during a working memory task as resting state connectivity. It also predicts that benefits on negative symptoms wane when the drug is stopped. However, it is possible glutamate actions could also be neuroprotective (see H1 above) whether or not it enhances glutamate function in the short-term.

Trial design

This is a 6-centre, double-blind, randomised placebo-controlled, efficacy and mechanistic study of minocycline added to standard APD treatment, versus standard APD treatment plus placebo for one year in patients in an acute episode of psychosis within 5 years of their first episode of psychosis. It will be carried out

in association with relevant Local Research Networks of the English National Institute for Health Research (NIHR) and Scottish Mental Health Research Networks (MHRNs). Project RAs (Clinical Research Officers: CROs) with the assistance of the local MHRN network Clinical Scientific Officers (CSOs) will recruit and assess patients in collaboration with local clinical teams and staff. A Research Manager (RM) will train and oversee the work of the CROs

Patients will be recruited while symptomatic within 5 years of onset when an inflammatory or other neurotoxic process may be active and susceptible to the various actions of minocycline. Consenting patients who meet inclusion criteria will be randomised to receive minocycline or matching placebo for one year, added to standard treatment organised by the clinical team. The progress of negative and other symptoms will be monitored at intervals through the year in parallel with a set of cytokine and imaging biomarkers and measures of social functioning. The stability of any changes after treatment is stopped will be assessed 3 months after the end of the trial period.

The trial will be monitored and managed by the PIs and the RM using the ethical, secure, research governance compliant and comprehensive project management procedures established by PsyGrid for the multi-centre study of first episode psychosis. These procedures are co-ordinated and automated using OpenCDMS software (www.opencdms.org), one of the major deliverables of PsyGrid. The OpenCDMS system organises on- or off-line collection of data onto a secure database, prompts for assessment, quality control, anonymisation of data, and randomisation of treatment allocation.

In this study patients will be allocated to treatment group according to a randomised permuted blocks algorithm, after stratification by centre. An experienced clinical trial pharmacist will oversee the blinding and unblinding procedures. Blinded supplies of placebo and active minocycline will be manufactured and distributed to local pharmacies by Catalent who have a long pedigree of involvement in clinical trials. Compliance will be assessed by interview and pill counts by the CROs or by the healthcare team at their monthly contacts.

The mechanistic biomarkers will probe specific hypotheses about how minocycline works to reduce negative symptoms and whether and how this is translated into improved social functioning.

Effective trial management is crucial. A trial management committee will meet monthly, chaired by the Chief Investigator. The trial will be overseen by an independent Trial Steering Committee (TSC) which meets six monthly, including patient/service-user representation. They will decide at the halfway point of recruitment on the basis of a blinded interim analysis whether it is ethical or feasible to continue the trial to achieve its stated objectives.

PLANNED INTERVENTIONS

Minocycline or matching placebo will be taken as 100mg tablets twice daily for 2 weeks, and increased 300mg daily for the remainder of the 12-month study period added to standard APD therapy and routine care. Catalent UK will organise the production of matching placebo and minocycline capsules, quality control, such assays as required by MHRA, labelling, distribution of supplies and dealing with returned medication.

PROPOSED OUTCOME VARIABLES

Primary clinical outcome variable:

1) Negative symptom severity as defined by negative syndrome subscale score on the . This is the gold standard for comprehensively rating symptoms of schizophrenia. The negative symptom subscale is composed of 7 items each rated 1-7. (19).

Primary biomarker outcome variables:

1) Medial prefrontal grey matter volume, (H1)
2) Circulating cytokine IL-6 concentration, (H2).
3) Dorsolateral prefrontal cortex blood oxygen level dependent (BOLD) response, % correct and connectivity during the N-back task, (H3)

Secondary clinical outcome variables:

1) Body weight and body mass index (BMI),

2) Full scale and positive syndrome subscale score ,

3) Functional outcome:

- Global Assessment of Function (GAF) from DSMIV
- Social Functioning Scale (SFS) self-rating in 7 domains
- Quality of Life Scale (QLS) for treatment effects related to deficit or negative symptoms

4) Cognitive outcome:

- Blyler WAISIII short form; current IQ;
- IQ decline from premorbid IQ (WTAR) predicts later negative symptoms.
- Digit-symbol test; processing speed
- Verbal fluency (VF). Cognitive correlate of negative symptoms

5) APD treatment in chlorpromazine equivalents

Side-effects and co-morbidity

1) Calgary Depression for Schizophrenia Scale (self- rating)

2) Extrapyramidal symptoms (EPS):

- The Simpson and Angus scale for pseudo-Parkinsonian symptoms and signs
- The Barnes Akathisia scale
- The Abnormal Involuntary Movements Scale for tardive dyskinesia (AIMS)

3) APD subjective side-effects: Antipsychotic Non-Neurological Side-Effects Rating Scale
(ANNSERS) developed in the CUtLASS study and used in PsyGrid.

4) 7-Point treatment adherence scale.

Almost all rating scales are in use in following-up PsyGrid patients and are loaded onto the OpenCDMS system.

Secondary biomarker outcome variables are changes in:

1) total and other regional grey matter volumes (H1),

2) cytokine screen (H2).

3) resting state connectivity (H3),

Exploratory analyses

Premorbid and current IQ, DUP and cytokine genotypes will be explored as predictor variables for response to minocycline.

FREQUENCY AND DURATION OF FOLLOW UP

Randomisation visit

This covers baseline ratings and 45mins scanning. These activities may be spread over more than one day. The CROs will check consent to continue and complete the items listed. Saliva will be collected using the Oragene kit which will be posted to the Manchester BRC biobanking facility. The patient will attend their local scanner unit for a 45 min scan. The session will be run by the radiographer and the CRO who will control computer projection of the N-back task and recording of performance.

The CROs will inform OpenCDMS that the patient should be randomised. OpenCDMS will: i) allocate the patient to a treatment arm; ii) email the local pharmacy about which treatment 'kit' to use to dispense two months supply; iii) inform the Manchester co-ordinating pharmacy who retain the coding key; and iv) notify Catalent that a local kit has been started. Patients will normally collect their own medication, but any convenient arrangement that improves compliance may be made.

Two, six and nine month visits

See the table of assessments. The CRO will receive emailed prompts for these visits from the CDMS. The CROs will arrange 2 monthly pill counts at the time the patients renew their trial treatment via the CRO, CPN or pharmacy. A second dipstick test for drugs of abuse will be carried out at 6 months and action taken according to results.

Twelve Month final trial visit

This will recap the screening visit safety measures and subsequent effectiveness measures. At all visits the patients cumulative clinical drug treatment will be updated from the case notes. The SCID interview will be repeated to ascertain diagnostic status. The scanning session will be repeated. Trial Medication will cease.

Fifteen month trial follow-up visit

Safety and efficacy measures will be repeated.

SAMPLE SIZE AND POWER CALCULATIONS

The study will be carried out in Trusts within 4 English MHRN hubs and 1 Scottish MHRN hub. The Northwest (NW) hub has been counted as two recruiting centres because it includes the Lancashire Care NHS Foundation Trust, one of the largest mental health Trusts in the country with a catchment area of 1.5 million covering several hospitals which are located at some distance from Trusts in the central Manchester conurbation. This gives six recruiting sites, 5 with a CRO and a sixth with the CRO/Research Manager. All NW MRI scans will be carried out in the University of Manchester in Salford which is accessible from central Manchester and Lancashire recruiting sites. There are therefore 5 imaging centres, 4 of which have harmonised procedures through their involvement with the NeuroPsyGrid consortium. In summary, there are 6 recruiting centres and 5 scanning centres nearly all of which have worked together as part of PsyGrid and NeuroPsyGrid.

The study is designed to produce clinical and biomarker data in 170 patients completing 12 months of placebo or minocycline add-on treatment (85 per group). This will ensure 90% power to detect a standardised effect size of 0.5 in the primary clinical outcome (e.g. a group difference in negative symptom scores of 3 units, assuming the within-group standard deviation is equal to 6 – as estimated from the Manchester-led MRI trial and PsyGrid clinical data) using a 2-tailed t-test at $p<0.05$. A difference of 3 units is the smallest effect we would consider to have any clinical significance. A simple t-test produces a conservative estimate. Power will be greater in practice using a repeated measures design and conditioning on

relevant baseline covariates. For statistical reasons we have chosen not to base our sample size calculations on mediator variables or on their hypothesised relationship with the primary outcome. However, based on the NeuroPsyGrid 5-site imaging data, the minimal detectable difference in grey matter is 2% at 80% probability with the sample size calculated above. This is much less than published MRI changes over one year.

We have based our calculations about recruitment on previous experience with the PsyGrid MHRN consortium and our previous minocycline study. In the Chaudhry et al study (2) in Pakistan and Brazil, 25% of those assessed were randomised 29% patients dropped out during the trial to an equal extent in both arms. We have assumed a 25% drop-out rate both from screening and from randomisation onwards. These figures are intended to be pessimistic and drop-out rates may be less than in Pakistan and Brazil because the proposed research is integrated with the clinical care of the patients and through demonstrating the involvement of patients and patient organisations in the design and monitoring of the study. The assumed drop-out rates give figures of 282 at screening and 226 at randomisation to produce completion in 170 (see appended flow chart). Each recruitment centre therefore needs to screen 2.1 a month and randomise 1.7 per month (see appended CRO workload calculations). Such figures were achieved by several centres in the PsyGrid study. The CRO workload is realistic. It allows a two-month training period and a maximum work rate of one combined MRI scanning and clinical rating session per week and 1-2 clinical follow-up ratings per week.

THE PROPOSED ANALYSES

All main analyses will be carried out at the end of the last follow-up assessments (i.e. there will be no interim analyses) and will be based on the intention-to-treat principle, with due consideration being given to potential biases arising from loss to follow-up.

All data analysis will be carried out using Stata (currently version 14) and Mplus (currently version 7).

DATA ANALYSIS PLAN – DATA DESCRIPTION

1. **Recruitment and representativeness of recruited patients**

Patient recruitment and flow through the trial will be described according to CONSORT Guidelines and illustrated using a 'CONSORT' flow chart. This will include monitoring of patterns of adherence to allocated treatment and patterns of attrition (missing data due to loss to follow-up).

2. **Baseline comparability of randomized groups**

Here baseline balance will be assessed through visual examination of descriptive statistics only (contingency tables for categorical variables; means, standard deviations and ranges for quantitative measures). There will be no statistical significance testing for baseline balance.

Baseline balance to be assessed for

Demographic and Clinical Variables: Centre; Age; Gender; Duration of Untreated Psychosis; Body Weight and Body Mass Index; IQ; Current Medication; Positive, Negative, General and Total Scores, Calgary Depression Scale, Global Assessment of Functioning; Social Functioning Scores; Quality of Life Score.

Biomarkers: Medial prefrontal grey matter volume; Circulating cytokine IL-6 concentration; Dorsolateral prefrontal cortex blood oxygen level dependent (BOLD) response, % correct and connectivity during the N-back task; cytokinine levels.

3. **Acceptance and adherence to allocated treatment**

This assessed by the Treatment compliance self-rating questionnaire completed by patients at each visit

4. Loss to follow-up and other missing data

The amount and patterns of missing outcome data will be tabulated. Their relationship with non-adherence, protocol violations and baseline characteristics of the participants will be described.

5. Descriptive statistics for outcome measures

At each follow-up time (2, 6, 9 and 12 months): as randomized (Inetion-To-Treat) and in terms of actual treatment received.

This will apply to all primary and secondary clinical outcomes, and to the primary and secondary biomarker measures.

As for the baseline assessments, tables will present means, standard deviations and ranges for all quantitative outcomes.

DATA ANALYSIS PLAN – FORMAL ANALYSIS

Treatment effect estimation

Treatment effects will reported using 95% confidence intervals, supplemented by their associated p-values. An effect will be regarded as statistically-significant if the p-value is less than or equal to 0.05, reduced to 0.01 if treatment effects at each of the four follow-up times are being evaluated separately.

Primary Outcome

Treatment effects on severity of negative symptoms will be estimated through the use of a random effects regression model (using Stata's xtreg command) after allowing for time of follow-up (2, 6, 9 or 12 months – treated as a categorical variable), Centre and baseline (pre-randomisation) severity of negative symptoms. The effect of time of follow-up on treatment efficacy will be evaluated by the treatment by time interactions and treatment effects at the four follow-up times estimated accordingly. If there is no significant variation in the treatment effect over

time (i.e. no treatment by time interaction) then the interaction will be dropped from the model and treatment efficacy common to all four follow-up times will be estimated. All models will contain centre by time and baseline severity by time interactions. Sensitivity of efficacy estimates to non-adherence and associated loss to follow-up will be assessed as described below.

Primary biomarker outcome variables

Similar methods to those used for the primary clinical outcomes will also be used here.

Secondary Outcomes

Similar methods to those used for the primary clinical outcomes will also be used here.

Non-adherence and missing outcome data

If there is a non-trivial amount of non-adherence to allocated medication, then treatment efficacy will be estimated through Complier-Average Causal Effect (CACE) estimation. Both ITT and CACE estimates will allow for missing data (drop outs), either through the use of inverse probability weights or the use of maximum likelihood estimation involving joint modeling of treatment effects on the outcome and of the missing data mechanisms. The presence of missing outcome data is likely to be highly-correlated with treatment non-adherence and non-adherence will be a key component of models generating inverse probability weights and of any explicit missing data model used in the likelihood-based methods. Note that this will be the case even when the main aim of the analysis is the estimation of ITT effects. In both the ITT and CACE approaches, group differences in outcomes and putative mediators will be evaluated using random effects models for longitudinal data (using Residual Maximum Likelihood, REML, as the fitting criterion), allowing for treatment centre and other baseline covariates. Tests of the mechanistic hypotheses (i.e. mediation), and their sensitivity of the results to possible confounding, will use structural equation modeling (including instrumental variable methods). Again, an important component of these analyses and the interpretation of the results will involve making sensible use of data on non-adherence and missing outcomes.

Construction of inverse probability weights

Treatment-effect estimates may be sensitive to assumptions concerning mechanisms of attrition. To investigate patterns of attrition, a logistic regression model will be used to explore what baseline characteristics, together with treatment allocation and patterns of adherence to allocated treatment, predict who will provide relevant outcome data up to each individual follow-up time, up to 12 weeks after randomization (separately for each randomized arm). If deemed necessary, the final models will be used to generate an expected probability of providing outcome data (for each arm separately) and the reciprocal of this estimated probability will be used as an inverse probability weight for use in the random effects models for the primary and secondary outcomes.

Analyses of Treatment Mechanisms (Treatment effect mediation)

At this stage of the analysis the strategy inevitably becomes more exploratory and, to some extent, the exact analysis will be dependent on the results of the main ITT analyses of both the primary clinical outcome and that of the mechanistic biomarkers. Structural equation models will be created to jointly model the treatment trajectories of both the primary clinical outcome and the biological mechanistic marker (or markers). Possibilities include parallel process models, latent change models and the possibility of confounding adjustments through the use of instrumental variables. A key component will be the inclusion of baseline measures of both the clinical outcome and the putative mediator, as these are likely to be the main sources of confounding of the effects of mediator on outcome.

Approaches to treatment effect mediation will be similar to those described in the published EME Report for the Worry Intervention Trial (WIT) – see Freeman *et al*. (2015) and the HTA Report on the statistical methodology relevant to the analyses of EME trials- Dunn *et al.* (2015).

References:

Dunn, G., Emsley, R., Liu, H., Landau, S., Green, J., White, I., &.Pickles, A. (2015). Evaluation and validation of social and psychological markers in randomised trials of complex interventions in mental health: a methodological research programme.

Health Technology Assessment **19**, Issue 93. DOI: http://dx.doi.org/10.3310/hta19930

Freeman, D., Dunn, G., Startup, H. & Kingdon, D. (2015). An explanatory randomised controlled trial testing the effects of targeting worry in patients with persistent persecutory delusions: the Worry Intervention Trial (WIT). *Efficacy and Mechanism Evaluation* **2**, Issue 1. DOI: http://dx.doi.org/10.3310/eme02010