Study Design "MOSAIC MS – MetabolOmic Signatures of disease Activity In predicting Clinical outcomes in Multiple Sclerosis"

Collaborators:

1. Dr Upcott [1], Dr Tallantyre[1,2], Dr Sam Loveless[2], Prof Robertson[1,2], Dr Kreft[3]

Statistical advisor:

Prof McKeever [3]

Affiliations:

1. Helen Durham Centre for Neuroinflammatory Disease, University Hospital of Wales, Heath Park, Cardiff, United Kingdom
2. Institute of Psychological Medicine and Clinical Neuroscience, School of Medicine, Cardiff University, Heath Park, Cardiff, United Kingdom
3. University of Nottingham, School of Medicine, Nottingham, UK

Background

Clinical relapse in Multiple Sclerosis (MS) represents a period of heightened inflammatory activity and are associated with disease progression. Despite advances in disease monitoring, there remains an unmet need for reliable, blood biomarkers capable of identifying active inflammatory disease and improving the detection of both clinical and subclinical disease activity.

The Welsh Neuroscience Research Tissue Bank has established a unique biorepository comprising cerebrospinal fluid (CSF) and plasma samples collected from people with MS. Untargeted whole-metabolome profiling has been performed on 90 CSF samples and 2,098 plasma samples of 835 MS patients using the Metabolon platform (443 CSF and 1338 plasma metabolites).

The study aims to identify metabolites associated with clinical relapse and radiological disease without clinical relapse activity in MS using untargeted metabolomic data. By characterising metabolic signatures associated with active inflammation, this study seeks to identify non-invasive candidate biomarkers of disease activity and improve understanding of the biochemical pathways underlying active inflammatory processes in MS.

Study Design

1. Using paired cerebrospinal fluid (CSF) and plasma samples (n=90), plasma metabolites will be stratified into three groups based on the strength of their correlation with the corresponding CSF metabolite. Metabolites will be classified as: (i) those demonstrating a strong CSF–plasma correlation (Spearman's correlation coefficient, $\rho > 0.5$), (ii) those demonstrating a weak or absent correlation ($\rho \leq 0.5$), and (iii) metabolites detected in plasma but not detectable in CSF.

2. Sensitivity analyses will then be undertaken to determine the effect of potential confounding variables, including age, sex, body mass index (BMI), and estimated glomerular filtration rate (eGFR), on the observed CSF–plasma metabolite correlations.

<u>Relapses and metabolites</u>

3. Plasma samples collected during a period of clinical relapse will be identified. A clinical relapse will be defined as the development of new or worsening neurological symptoms persisting for more than 24 hours in the absence of infection, fever, recent history of vaccination or another alternative explanation.

Relapse events will have been documented and confirmed prospectively by a clinician within the South Wales Neuroinflammatory Service.

4. Participants with a clinical relapse sample will be stratified according to the availability of longitudinal samples collected before and/or after the relapse event. Patients will be classified into one of three groups: (i) those with pre-relapse, relapse, and post-relapse samples (ii) those with pre-relapse and relapse samples only and (iii) those with relapse and post-relapse samples only. This stratification will facilitate longitudinal analyses of metabolite changes occurring before, during, and following clinical relapse. A plasma sample of during a relapse needs will be defined as a sample to be taken within 42 days of the onset of symptoms and before commencement of corticosteroid treatment.

5. Primary outcome is changes in metabolite expression during a relapse in participants with paired pre-relapse, relapse, and post-relapse plasma samples (n=46), linear mixed-effects modelling will be used to identify metabolites associated with clinical relapse. Models will include participant-specific random effects to account for variation in baseline metabolite expression, repeated measures and will be adjusted for age, sex, body mass index (BMI) and estimated glomerular filtration rate (eGFR) as potential confounding variables.

6. For internal replication, linear mixed-effects modelling, adjusted for age, sex, body mass index (BMI), and estimated glomerular filtration rate (eGFR) will then be applied to the two remaining longitudinal cohorts: participants with paired pre-relapse and relapse samples (n=45), and participants with paired relapse and post-relapse sample (n=29). Metabolites statistically significant (nominal $p<0.05$) in both step 5 and 6 and with a directionality concordant, the effect size will be considered clinically relevant. Given the independent replication in two additional cohorts, adjustment for multiple testing is not considered necessary.

7. Secondary sensitivity analysis: plasma samples from patients with a confirmed clinical relapse according to the criteria outlined in point 3 will be combined with samples not taken during a relapse from an independent cohort of patients matched for age and sex, and random forest classification model will be constructed. The model will identify a metabolite panel to discriminate between relapse and stable disease state.

<u>MRI activity without clinical relapse</u>

8. Plasma samples collected before and after (2-year interval to either side of) an MRI, during a period of clinical stability with no intervening relapse. This interval is chosen to limit the effect of biological ageing on metabolite expression.

Subsequently, two groups will be compared, those patients with a stable MRI (reference group, n=27) and patients who developed new lesions compared to the previous MRI (n=33). These samples will be used to identify metabolites associated with new radiological disease activity independent of clinical relapses. Using these longitudinal plasma samples, linear mixed-effects models will again be employed, the model will include an interaction term metabolite*MRI status, to compare whether the delta of metabolite between baseline and follow up sample is different between the radiologically stable and active patients.

9. Secondary sensitivity analysis will include (i) Spearman's rho correlation analysis between the number of new T2 lesions and the delta in metabolites between sample 1 and 2, and directionality concordance will be compared with the results of step 8. (ii) Random forest classification model will be constructed comparing the change in metabolite expression between sample 1 and 2 and MRI status (stable vs active). Given the stringent inclusion criteria and the associated relatively low number of patients in the analysis, FDR adjustment for multiple testing is not deemed feasible. Metabolites statistically significant (nominal $p<0.05$) in both step 8 and 9 ($p<0.05$ for Spearman's rho with a directionality concordant effect size or high importance score in random forest results) will be considered clinically relevant.

As all analyses are designed to assess the effect of disease activity (clinically or radiologically) on metabolite expression, disease modifying treatment effects will not be taken into account in the modelling (e.g. if a patient is on a DMT and has disease activity, this indicates breakthrough disease = active disease).

Depending on the number of the metabolites identified, pathway enrichment analysis will be considered to investigate the metabolic pathways associated with clinical relapse and radiological disease activity to further understand the underlying biology.