A UK Biobank-Based Metadata And Multi-Omics Analysis Approach Exploring Comorbidities And Biomarkers In Complex Disorders Like Multiple Sclerosis

Jaya Singh, Vaishali Chakraborty, Jemima Jacob, Manuela Samuel, Ashish Kumar Choudhary
Strand Life Sciences, Bangalore, India

Contact
Jaya Singh
AVP -
Business Development
+91 96320 03280
jaya@strandls.com
VISIT **LINKEDIN**

strand

VISIT OUR **WEBSITE**

Introduction

- Multiple sclerosis (MS) is a chronic demyelinating disorder affecting ~2 million people globally.
- Comorbidities can worsen outcomes, accelerating disability, reducing life quality, and increasing mortality. The significance of following a healthy lifestyle on comorbidities with a potential impact on MS is less studied.
- We analyzed metadata from MS and non-MS individuals with healthy lifestyles and explored potential MS biomarkers using pQTL analysis of Olink proteomics data from the UK Biobank.

Objective

- **Cardiometabolic Comorbidities in MS:** Assess associations of hypertension (HT) and hyperlipidemia (HL) with MS in adults who otherwise follow healthy lifestyle practices.
- **pQTL Analysis and Biomarker Discovery:** Leverage pQTL analysis to identify SNPs influencing protein expression in MS vs. non-MS individuals and evaluate their biomarker potential.
- **Understanding Disease Mechanisms:** Map protein expression profiles to immune pathways, neuroinflammation, and demyelination processes underlying MS.

Methods

Step	Description	Details
1 **Data**	Source	UK Biobank (MS-2123, Controls - 500,496)
2 **Cohorts**	Metadata cohort	MS defined by ICD-10 code G35. Restricted to healthy-lifestyle participants (non/previous smokers, BMI <30kg/m2, regular exercise). HT & HL as covariates. **Cohort size:** MS - 265, Control - 99,112.
	Proteomics cohort	MS cases vs. non-MS controls with Olink proteomics data.
		Cohort size: MS = 436, Control = 3003
3 **Metadata analysis**	Logistic regression	Tested association of MS with HT, HL, age, sex, smoking, alcohol, and HT×HL interaction
4 **Proteomics**	Differential protein	QC → design matrix → limma. FDR < 0.05 cutoff
	pQTL mapping	Regenie two-step pipeline:
		- Step 1: Polygenic model
		- Step 2: Per variant association
		QC: allele count ≥10, MAF ≥0.01, missing genotype rates ≤10%, HWE p >1e-15
		Significance: genome-wide $p < 5x10^{-8}$
5 **Downstream analysis**	Variant annotation	Ensembl Variant Effect Predictor (VEP)
	Pathway enrichment	ClusterProfiler & Signaling Pathway Impact Analysis (SPIA)

Highlights

- **Hyperlipidemia** was associated with a **2.2x higher risk of MS disease progression** in participants aged 40-55 yrs.
- **Males** were associated with **a lower risk for MS**.
- **189 proteins** showed **significantly altered** expression levels: 139 upregulated and 50 downregulated in MS vs. controls.
- **The key markers - CXCL13, FABP4, GDF15, CCL7** and **CXCL17**, have published evidence in the pathogenesis, prognosis and therapy response of MS.
- Pathway analysis emphasized immune dysregulation, with **significant activation of the chemokine and cytokine signaling pathways in MS**.
- UK Biobank-enabled integrative data analysis demonstrates the value of combining large-scale phenotype and omics data, and reveal the multifactorial nature of MS involving immune cell trafficking, chronic inflammation, and metabolic dysregulation.

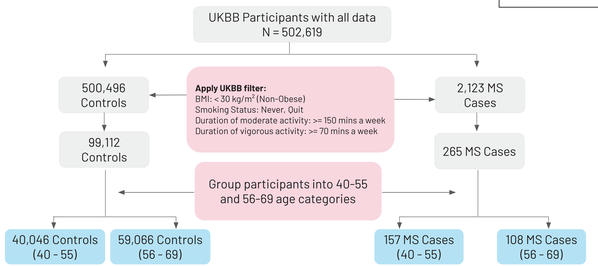

Figure 01: Study design and participant selection from UK Biobank workflow

Results and Conclusions

Table 01: Association of comorbidities and lifestyle factors with multiple sclerosis risk

Factor	Effect on MS	Statistical outcome
HL (40-55 yrs)	↑ risk, 2.2x	OR ≈ 2.20, p = 0.007
HT	Not significant	OR ≈ 0.88, p = 0.55
HT × HL	Not significant	OR ≈ 1.46, p = 0.53, Independent effects
Male	↓ risk	OR ≈ 0.46, p = 5.12e-05
Previous smoking	↑ risk	OR ≈ 1.38, p = 0.047

Table 02: Protein biomarkers and their associations with multiple sclerosis, highlighting prognostic, pathogenic, and potentially protective roles

Protein	Association in MS	Interpretation
CXCL13	Prognostic marker; linked to B cell-driven inflammation and disease activity	Predicts aggressive disease/ progression; associated with favorable treatment outcomes
FABP4	Increased disability scores (EDSS)	Especially in pediatric MS and secondary progressive MS
GDF15	Elevated in stable MS	Suggests compensatory, tissue-protective or anti-inflammatory response
CCL7	Significantly upregulated	Promotes immune cell recruitment and lesion development
CXCL17	Appears protective	Modulates immune responses, may limit disease severity

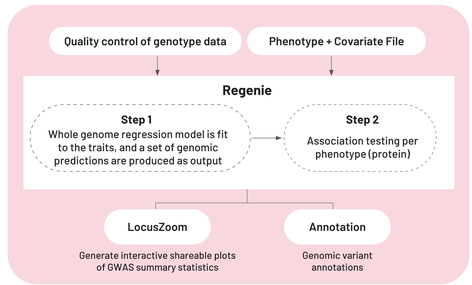

Figure 02: pQTL analysis workflow using Regenie

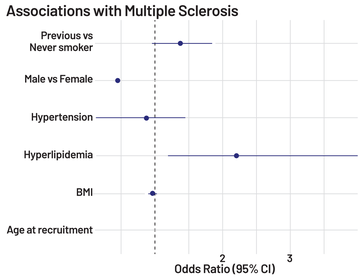

Figure 03: Association of comorbidities and lifestyle factors with multiple sclerosis risk

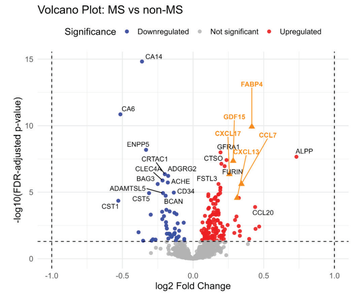

Figure 04: The volcano plot showing the differentially expressed proteins between MS and non-MS

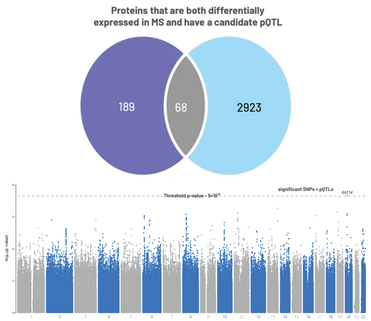

Figure 05: Manhattan plot showing genome-wide association results for pQTLs, with significant SNPs surpassing the threshold ($p < 5\times10^{-8}$), including a notable locus at KAT14