medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Low lymphocyte count is a risk factor for Parkinson's disease

Melanie P Jensen (MBBS)*[1], Benjamin Meir Jacobs (MRCP)*[1], Ruth Dobson (MRCP, PhD)[1], Sara Bandres-Ciga (PhD)[2], Cornelis Blauwendraat (PhD)[2], Anette Schrag (MD, PhD)[3], The International Parkinson's Disease Genomics Consortium (IPDGC), Alastair J Noyce (MRCP, PhD)[+1,3]

* equal contribution

Affiliations

1. Preventive Neurology Unit, Wolfson Institute of Preventive Medicine, Queen Mary University of London, London, UK
2. Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA
3. Department of Clinical and Movement Neurosciences, UCL Institute of Neurology, London, UK

Corresponding author ([+])

Dr Alastair Noyce

Preventive Neurology Unit, Wolfson Institute of Preventive Medicine, Queen Mary University of London, Charterhouse Square, London, EC1M 6BQ

Email - a.noyce@qmul.ac.uk Telephone - +44 (0) 207 882 5841

Word count: 2,991

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Key points

Question: Is the leukcoyte differential count a feature of pre-diagnostic Parkinson's disease?

Findings: In the UK Biobank, a longitudinal cohort study with over 500,000 participants, lower lymphocyte count was associated with a 23% increased odds of incident PD, a significant difference. Mendelian randomisation revealed a convincing causal effect for low lymphocyte count on PD risk.

Meaning: Pre-diagnostic Parkinson's disease is associated with lower lymphocyte counts; the suggestion of causal effect may shed light on PD pathogenesis.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Importance: Biomarkers for the early detection of Parkinson's disease (PD) are needed. Patients with PD display differences in peripheral blood biomarkers of immune function, including leukocyte differential counts and C-reactive protein (CRP), compared to controls. These differences may be useful biomarkers to predict PD, and may shed light on PD pathogenesis.

Objectives: To identify whether peripheral immune dysregulation is a pre-diagnostic feature of PD, and whether it plays a causal role.

Design: Cross-sectional association analysis of the relationship between differential leukocyte count and other markers of acute inflammation at enrolment, and incident cases of PD in UK Biobank. We used Mendelian randomization to establish whether differences in leukocyte differential counts have a causal influence on risk of PD.

Setting: UK Biobank; a population-based cohort with over 500,000 participants aged 40-69 recruited in the UK between 2006 and 2010.

Participants: PD cases were defined as individuals with an ICD-10 coded diagnosis of PD. Cases were defined as 'incident' if their age at diagnosis was greater than their age at recruitment to UKB. 'Controls' were defined as individuals without a diagnosis of PD. After applying exclusion criteria for pre-existing health conditions that can influence blood counts, 507 incident PD cases and 328,280 controls were included in the analysis.

Exposure: Blood cell markers (absolute and relative counts) and other markers of inflammation were obtained from blood tests of participants taken at the initial visit.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Results: Lower lymphocyte count was associated with increased odds of incident PD (odds ratio [OR] 0.77, 95% confidence interval [CI] 0.65-0.90). There was weaker evidence of association between lower eosinophil and monocyte counts, lower CRP, and higher neutrophil counts on risk of incident PD. The association between lymphopenia and incident PD remained robust to sensitivity analyses. Mendelian randomization analyses suggested that the effect of low lymphocyte count on PD risk was causal (OR 0.91, 95% CI 0.85 - 0.99).

Conclusions and relevance: In this large, prospective setting, lower lymphocyte count was associated with higher risk of subsequent PD diagnosis. Furthermore genetic evidence supported a causal role for lymphocyte count on PD risk.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

Parkinson's disease (PD) affects 2% of the population over 65.[1] The diagnosis is made once motor signs appear, however by this stage ~50% of nigrostriatal neurons have been lost.[2] There is an urgent unmet clinical need for earlier identification of PD and development of therapies which could slow, prevent, or reverse the progression of the disease.

Immune dysregulation may play a key role in the pathogenesis of PD. The white blood cell (WBC) differential is a crude marker of immune function but is amenable to explore in large-scale observational studies. Studies reveal lower lymphocyte counts in PD patients compared with controls, driven by reductions in helper-CD4$^+$, rather than cytotoxic-CD8$^+$, T-cell and B-cells counts; hypothesized to represent a cytotoxic immune signature.[3-7] Case-control studies have also identified higher neutrophil and lower lymphocyte counts in patients with established PD compared with controls.[8]

Genetic, epidemiological, and cytokine profiling studies have refined this area of study.[9] Human Leukocyte Antigen variants (HLA-DRB1/DRB5) have been identified as risk loci for PD in genome-wide association studies (GWAS).[10,11] Large-scale polygenic risk score analyses suggest this pathway contributes to PD heritability.[12] *In vitro*, alpha-synuclein-derived peptides are preferentially displayed on major histocompatibility (MHC) molecules associated with PD risk, driving proinflammatory T-cell responses.[13,14] Variants in the leucine-rich repeat kinase 2 (*LRRK2*) gene, a target for proinflammatory signals, confer effects in the same direction on risk for PD and Crohn's disease, suggesting common genetic links.[15] Observational studies have reported reduced risk of PD and reduced penetrance in *LRRK2*-associated PD with use of immunosuppressants and non-steroidal anti-inflammatory drugs.[16,17] The prospective ICICLE-PD cohort study found that a baseline 'pro-inflammatory'

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

cytokine serum profile in PD patients was associated with faster motor deterioration than an 'anti-inflammatory' profile.[18]

Whether immune dysregulation occurs as an early feature of PD and may be a source of biomarkers, or immune dysregulation has a causal role in driving PD initiation and progression, remains unclear. We studied the relationship between differential leukocyte count and biochemical markers of acute inflammation at enrolment, and incident cases of PD in the UK Biobank (UKB) (https://www.ukbiobank.ac.uk), a large longitudinal cohort with ~500,000 participants. We followed-up signals detected to determine whether differences in leukocyte subsets exerted a causal influence on PD risk using Mendelian randomization (MR).

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Methods

Population

UKB recruited ~500,000 individuals aged 40-69 between 2006-2010; prospective follow-up data, including census data, blood tests, and healthcare records, are regularly obtained.[19]

PD cases were defined as individuals with an ICD-10 diagnosis of PD (code G20) derived from Hospital Episode Statistics or a self-reported diagnosis of PD. Date at PD diagnosis was determined using the UKB data field 'Date of Parkinson's Disease report'. Age at diagnosis was derived using this field, age at recruitment, and birth year. Cases were defined as 'incident' if their age at diagnosis was greater than at recruitment. 'Prevalent' PD cases, i.e. with a diagnosis of PD at baseline, were excluded from analyses. 'Controls' were defined as all other individuals in the dataset after applying these exclusions.

Various pre-existing health conditions can influence blood counts. To minimize bias from unbalanced comorbidities among cases and controls we excluded individuals with ICD-10 diagnoses of malignant neoplasms, disease of the blood and blood-forming organs, autoimmune disease, thyrotoxicosis, demyelinating disease of the central nervous system, inflammatory respiratory conditions (asthma and bronchiectasis), non-infective enteritis, inflammatory dermatological conditions (atopic dermatitis and psoriasis), inflammatory polyarthropathies, spondylopathies, and eating disorders (supplementary table 1).

Blood cell markers

Blood cell markers (absolute and relative counts) and other markers of inflammation (CRP and albumin) were obtained from baseline blood tests of UKB participants taken at the initial

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

assessment visit. Details of data processing can be found on the UKB website (http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/haematology.pdf).

Statistical analysis

We determined associations of blood cell and inflammatory markers with incident risk of PD using logistic regression. As our primary analysis, we conducted multivariable logistic regression, modelling incident PD diagnosis as the outcome and adjusting for age, sex, Tonwsend deprivation score, and ethnicity (dichotomized as 'White' background vs all other ethnicities). Models were of the form: Incident PD~Age+Sex+Deprivation+Ethnicity+blood cell marker. The strength of association was determined using the likelihood ratio test, comparing the full model to a null model consisting of the confounding covariates only (Incident PD~Age+Sex+Deprivation+Ethnicity).

We then undertook a variety of sensitivity analyses. First, we included additional covariates in the models: body mass index (BMI) at recruitment, smoking status ("ever" vs "never"), and alcohol consumption ("ever" vs "never"). Second, we excluded individuals within serial time windows of PD diagnosis (<1, <2, <3 years from diagnosis etc.) to determine whether the effects from the primary analysis were restricted to individuals who would go on to develop PD sooner. Third, we repeated the analysis in a matched case:control analysis, individually matching controls by age and sex to PD cases 4:1.

To determine associations between blood markers and time until PD diagnosis, we constructed linear models for the inverse-normal-transformed time to PD diagnosis on age, sex, Townsend score, ethnicity, and blood cell marker. Model fit was quantified using the likelihood ratio test.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Mendelian randomization

We used MR to determine whether the observational association between lower lymphocyte count and incident PD represents a causal relationship. For MR, SNPs associated with the exposure of interest are used as instrumental variables to estimate the causal effect of the exposure on the outcome.[20,21] The random allocation of genetic variants from parent to offspring and lifelong exposure mean there are advantages over traditional observational studies in reducing confounding and reverse causation.[22] We performed two-sample MR using the TwoSampleMR R package.[23,24] For the exposure instrument, we used summary statistics from the largest published GWAS on blood cell traits from ftp://ftp.sanger.ac.uk/pub/project/humgen/summary_statistics/UKBB_blood_cell_traits/.[25] This GWAS of 408,112 European UK Biobank participants used as its outcome measure the absolute lymphocyte count adjusted by confounding covariates (age, sex, principal components, study-specific factors), and rank inverse-normalised. Thus a one unit increase in the beta coefficient represents a 1 standard deviation increase in adjusted and normalised lymphocyte count per additional effect allele.

We applied the following to develop a genetic instrument for lymphocyte count:

1. Removed SNPs not typed/imputed in the outcome GWAS;
2. Restricted to biallelic single nucleotide variants;
3. Removed SNPs within the super-extended MHC region due to the complex pleiotropy of this region (hg19, chr6:25,000,000-35,000,000);
4. Restricted to SNPs strongly associated with standardized lymphocyte count (p-$value_{adjusted}$<5e-08);
5. Clumped SNPs using stringent default parameters (LD window=10,000 kb, r^2=0.001).

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

6. Removed SNPs explaining more variance in the outcome than the exposure through Steiger filtering.[26]

Outcome data for PD were taken from the most recent and largest case-control GWAS of PD published by the IPDGC and 23andMe.[27] Overlap in the controls in the exposure and outcome data can result in bias, so we used summary statistics from the PD GWAS that excluded participants from UKB.

We harmonized exposure and outcome SNPs to ensure effect estimates were aligned for the same effect allele. As our primary analysis, we used the inverse-variance weighted (IVW) MR estimate, which provides an accurate causal estimate when MR assumptions are valid.[28] As secondary sensitivity analyses, we applied the Mixture of Experts approach, which applies different MR estimators and methods for SNP instrument selection (heterogeneity and directionality filtering), and predicts which method has the highest probability of accurately estimating the true causal effect based on the data characteristics.[23] Power calculations were performed using Stephen Burgess' online calculator (https://sb452.shinyapps.io/power/).

Data and code availability

UK Biobank data are available via application (https://www.ukbiobank.ac.uk/). Code is available at https://github.com/benjacobs123456/PD_FBC_UKB. PD GWAS summary statistics which exclude UKB are from Nalls *et al.* 2019 and an application to 23andMe https://research.23andme.com/dataset-access/.[27] Blood cell trait GWAS summary statistics have been made publicly available by the authors at ftp://ftp.sanger.ac.uk/pub/project/humgen/summary_statistics/UKBB_blood_cell_traits/.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Results

Demographics

After applying the exclusion criteria, 507 incident PD cases and 328,280 controls remained in the main (unmatched) analysis. Participant demographic data are shown in table 1.

Association of blood cell and inflammatory traits with incident PD

In a multivariable logistic regression model adjusting for age, sex, deprivation score, and ethnicity, there was evidence of association (false discovery rate [FDR] Q<0.05) between lower lymphocyte count and incident PD (odds ratio [OR] 0.77, 95% confidence interval [CI] 0.65-0.90, table 2, figure 1). There was evidence of association (FDR Q<0.05) between lower eosinophil count and incident PD but with wide CIs (OR 0.33, 95% CI 0.14-0.76, table 2). There was weaker evidence (FDR Q<0.10) of associations between lower CRP, lower monocyte count, higher neutrophil count and increased risk of incident PD (table 2).

To determine whether these associations could be driven by confounding, we constructed models in which we also controlled for variables which can impact both PD risk and blood cell indices: BMI, smoking, and alcohol consumption. The effect estimates from these models were less precise but of a similar magnitude to the primary analysis (lymphocyte count OR 0.78, 95% CI 0.66-0.92; eosinophil count OR 0.38, 95% CI 0.16-0.87; CRP OR 0.97, 95% CI 0.94-1.01; monocyte count OR 0.60, 95% CI 0.34-1.06; neutrophil account OR 1.08, 95% CI 1.01-1.15 on risk of incident PD) (supplementary table 2).

To examine the possibility that reverse causation could be driving our findings - i.e. that early PD could be driving lower lymphocyte counts - we excluded individuals who underwent blood draw within 5 years of PD diagnosis. Despite the loss of power resulting from this

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

restriction, we obtained similar effect estimates (supplementary table 3). To examine whether the association between lower lymphocyte count and higher PD risk persisted through the pre-diagnostic period, we repeated this analysis excluding individuals within serial time windows of diagnosis (within 1, 2, 3, 4, 5, 6, 7, and 8 years of diagnosis). Despite lower numbers of cases with increasing time pre-diagnosis, there was a consistent signal in all groups (supplementary table 4). We obtained similar results for the association between lower lymphocyte count and risk of incident PD with the matched case-control analysis (OR 0.78, 95% CI 0.65-0.93) (supplementary table 5). Lymphopenia considered as a binary trait (defined as absolute lymphocyte count<$1x10^9$ cells/L), was strongly associated with a higher risk of PD (OR 1.93, 95% CI 1.26-2.97, p=0.006). Exclusion of extreme lymphocyte counts (mean ± >3SD, leaving an inclusion lymphocyte count range 0.05-$3.85x10^9$ cells/L) did not substantially alter the observed association (OR 0.76 per 1-SD increase in lymphocyte count, 95% CI 0.64-0.90).

To assess whether inflammatory and blood cell markers were associated with time until PD diagnosis, we constructed linear models adjusting for age, sex, deprivation score and ethnicity. There was no strong evidence that any blood markers were associated with time until PD diagnosis (FDR >0.05, supplementary table 6).

Mendelian randomization

To generate an instrument for lymphocyte count, we excluded SNPs not typed/imputed in both the exposure and outcome datasets, SNPs with $p \geq 5e^{-08}$, SNPs lying within the super-extended MHC (chr6:25,000,000-35,000,000 on genome build hg19), and palindromic SNPs with intermediate effect allele frequencies. We performed LD clumping using default

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

parameters and Steiger directionality filtering, yielding a genetic instrument of 510 independent non-MHC autosomal SNPs associated with lymphocyte count (supplementary table 7). Collectively these SNPs explained 5.37% of the variance in lymphocyte count in this sample and could be considered a strong instrument (F statistic=45.3).

The primary MR analysis (IVW) showed evidence of a causal effect of lymphocyte count on PD risk commensurate with the size of the observational effect (OR 0.91, 95% CI 0.85-0.99, p=0.023). There was no evidence that unbalanced horizontal pleiotropy (whereby variants influence the outcome via pathways other than through the exposure) was biasing the IVW result (MR-Egger intercept 0.0006, p=0.76). There was evidence of substantial heterogeneity in the IVW estimate (Cochran's Q=642, $p=5.29e^{-05}$), however heterogeneity filtering did not alter the magnitude of the effect (OR 0.91, 95% CI 0.85-0.97, p=0.0034), and in fact increased the precision of the estimate. Standard sensitivity analyses and the 'mixture of experts' approach yielded similar causal effect estimates with varying degrees of precision, collectively providing evidence of a causal effect of genetically-determined lymphocyte count on PD risk (figure 2, supplementary table 8).

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Discussion

In this prospective cohort involving >320 000 individuals, lower lymphocyte count was associated with increased risk of incident PD; marginal associations were observed for lower eosinophil and monocyte counts, and higher neutrophil counts. The association between lymphopenia and incident PD risk remained robust to several sensitivity analyses. Follow-up MR analyses revealed evidence supporting the hypothesis that lower lymphocyte count may be a causal risk factor for developing PD later in life.

Only one study has explored the link between leukocyte subsets and risk of incident PD. In the Swedish Apolipoprotein-Related Mortality Risk cohort, lower lymphocyte count was associated with a lower risk of incident PD (HR 0.74, 95% CI 0.59-0.94).[29] However, pre-existing health conditions which can influence blood counts were not excluded or adjusted for so the possibility of residual confounding remained; we applied stringent exclusion criteria capturing comorbidities which influence leukocyte counts, such as autoimmune disease and cancer.

Several studies have assessed changes in leukocyte populations in PD after diagnosis. Established PD is associated with lymphopenia, driven by absolute reductions in $CD4^+$ T-helper cells, $CD19^+$ B-cells and Treg cells.[4,5,30,31] In a study of 123 newly diagnosed PD patients, the percentage of neutrophils and lymphocytes had positive and negative correlations, respectively, with UPDRS motor scores.[32] This has been interpreted as evidence to support an autoimmune/inflammatory component in PD pathogenesis.[33]

Reductions in the absolute number of naïve-T/B-cells is also a feature of age-related immune dysregulation, leading to the hypothesis that PD is a state of premature ageing.[34-36] Flow

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

cytometric analysis has demonstrated that leukocyte apoptosis is higher in PD patients than controls and is associated with dopaminergic deficits on SPECT.[37] However, studies to date have examined immune markers in established PD and cannot determine whether dysregulated adaptive immunity causally influences PD risk or represents a downstream consequence of PD pathology. Moreover, most PD cases in these studies are receiving dopaminergic medication, which could confound immunophenotypic patient/control differences.[38-40] We demonstrate that lower lymphocyte count may be a feature of PD at least 8 years before diagnosis, before the initiation of dopaminergic therapy, extending the observations made in patients with established PD into the pre-diagnostic phase.

The observed association between lymphopenia and incident PD may be driven by a causal effect, residual confounding or reverse causation. Our MR analyses suggest a causative effect of low lymphocyte count on PD risk. Specifically, the IVW estimate suggests a weak causal effect of lymphocyte count on PD risk such that each 1-SD increase in adjusted lymphocyte count decreased the odds of PD by 9%. The MR-Egger regression intercept, which quantifies the degree to which net unbalanced horizontal pleiotropy may bias the IVW estimate, was close to null, suggesting that the IVW estimate provides a robust estimate of the true causal effect. Sensitivity analyses using a variety of MR estimators and heterogeneity filtering supported the magnitude of the IVW estimate, albeit with varying precision. These data lend weight to the view that low lymphocyte count is not purely an effect of PD and its prodromes (reverse causation) or confounding, but may be involved in a causal chain of events contributing to the development of PD.

To address confounding in the observational study, we corrected for potential confounders in the primary analysis and undertook sensitivity analyses (further addition of covariates and a

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

matched analysis). However residual confounding, that is pre-diagnostic PD affects lymphocyte count, may remain. Although the effect persisted after excluding individuals who underwent blood draw within 8 years of diagnosis, pre-diagnostic features of PD including constipation may predate clinical diagnosis by 20 years.[41,42] The gut has been proposed as the site of initiation of PD and fecal microbiome changes are also noted in prodromal PD.[43] In turn, manipulation of intestinal microbiota in mice has been shown to directly modulate lymphopenia-induced autoimmunity.[44] Lymphopenia may be a feature of pre-diagnostic PD, and reflect early pathological changes in peripheral tissues.

Higher CRP was marginally associated with reduced risk of PD. A meta-analysis of 23 case-control studies found significantly higher levels of CRP in the serum, cerebrospinal (CSF), and whole blood of manifest PD subjects compared with controls.[45] CSF CRP levels tend to increase over time from PD diagnosis and the development of PD-dementia.[46,47] These findings suggest that systemic inflammation impacts or occurs as a consequence of disease progression, but does not shed light on whether it influences disease initiation.[9]

Strengths of our study include the large sample size derived from UKB. Despite low response rates potentially leading to more favorable risk factor profiles, exposure-outcome associations in the UKB seem to be generalisable.[48] The comprehensive phenotyping of individuals in the cohort allowed us to correct for multiple potential confounding factors. In contrast to previous work utilizing cohorts with manifest PD, the availability of baseline blood tests and longitudinal assessment of PD diagnosis enabled us to examine the association between leukocyte subsets and risk of incident PD.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Limitations of our study include the lack of adjustment for medication which could impact leukocyte subsets, although exclusion for comorbidity will have captured a significant proportion of this confounding. Due to lack of flow cytometric data in UKB, we could not establish whether the observed association was driven by reductions in T-cells and/or B-cells. MR analysis precludes the identification of non-linear exposure-outcome associations, although non-linear mechanisms are not seen in other conditions in which lymphopenia influences outcome.[49,50]

In conclusion we report that lower lymphocyte count was associated with higher risk of subsequent diagnosis of PD in a large UK cohort. The association remained robust to a range of sensitivity analyses and persisted at 8 years pre-diagnosis. MR analyses suggested that this relationship may be causal and this in particular warrants further exploration. It is also plausible that lymphopenia may be a consequence of pre-diagnostic PD which, although lacking specificity in isolation, could enhance efforts to identify those in the earliest stages of PD.[51] Further work is required to replicate these findings in other cohorts, address the mechanisms by which lymphopenia may causally intersect with the pathobiology of PD, and evaluate whether lymphopenia improves algorithms aiming to predict who will develop PD in prospective cohort designs.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Acknowledgements

This research has been conducted using the UK Biobank Resource. We thank the participants of the UK Biobank. We would also like to acknowledge the 23andMe Research Team (especially Dr Karl Heilbron) and the 23andMe research participants.

We acknowledge the IPDGC consortium members and affiliations:

United Kingdom: Alastair J Noyce (Preventive Neurology Unit, Wolfson Institute of Preventive Medicine, QMUL, London, UK and Department of Molecular Neuroscience, UCL, London, UK), Rauan Kaiyrzhanov (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Ben Middlehurst (Institute of Translational Medicine, University of Liverpool, Liverpool, UK), Demis A Kia (UCL Genetics Institute; and Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Manuela Tan (Department of Clinical Neuroscience, University College London, London, UK), Henry Houlden (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Catherine S Storm (Department of Clinical and Movement Neurosciences, UCL Queen Square Institute of Neurology, London, UK), Huw R Morris (Department of Clinical Neuroscience, University College London, London, UK), Helene Plun-Favreau (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Peter Holmans (Biostatistics & Bioinformatics Unit, Institute of Psychological Medicine and Clinical Neuroscience, MRC Centre for Neuropsychiatric Genetics & Genomics, Cardiff, UK), John Hardy (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Daniah Trabzuni (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK; Department of Genetics, King Faisal Specialist Hospital and Research Centre, Riyadh, 11211 Saudi Arabia), John Quinn (Institute of Translational Medicine, University of

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Liverpool, Liverpool, UK), Vivien Bubb (Institute of Translational Medicine, University of Liverpool, Liverpool, UK), Kin Y Mok (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Kerri J. Kinghorn (Institute of Healthy Ageing, Research Department of Genetics, Evolution and Environment, University College London, London, UK), Nicholas W Wood (UCL Genetics Institute; and Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Patrick Lewis (University of Reading, Reading, UK), Sebastian R Schreglmann (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Ruth Lovering (University College London, London, UK), Lea R'Bibo (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Claudia Manzoni (University of Reading, Reading, UK), Mie Rizig (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Mina Ryten (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Sebastian Guelfi (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Valentina Escott-Price (MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University School of Medicine, Cardiff, UK), Viorica Chelban (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Thomas Foltynie (UCL Institute of Neurology, London, UK), Nigel Williams (MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff, UK), Karen E. Morrison (Faculty of Medicine, University of Southampton, UK), Carl Clarke (University of Birmingham, Birmingham, UK and Sandwell and West Birmingham Hospitals NHS Trust, Birmingham, UK), Kirsten Harvey (UCL School of Pharmacy, UK), Benjamin M Jacobs (Preventive Neurology Unit, Wolfson Institute of Preventive Medicine, QMUL, London, UK).

France: Alexis Brice (Institut du Cerveau et de la Moelle épinière, ICM, Inserm U 1127, CNRS, UMR 7225, Sorbonne Universités, UPMC University Paris 06, UMR S 1127, AP-HP,

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Pitié-Salpêtrière Hospital, Paris, France), Fabrice Danjou (Institut du Cerveau et de la Moelle épinière, ICM, Inserm U 1127, CNRS, UMR 7225, Sorbonne Universités, UPMC University Paris 06, UMR S 1127, AP-HP, Pitié-Salpêtrière Hospital, Paris, France), Suzanne Lesage (Institut du Cerveau et de la Moelle épinière, ICM, Inserm U 1127, CNRS, UMR 7225, Sorbonne Universités, UPMC University Paris 06, UMR S 1127, AP-HP, Pitié-Salpêtrière Hospital, Paris, France), Jean-Christophe Corvol (Institut du Cerveau et de la Moelle épinière, ICM, Inserm U 1127, CNRS, UMR 7225, Sorbonne Universités, UPMC University Paris 06, UMR S 1127, Centre d'Investigation Clinique Pitié Neurosciences CIC-1422, AP-HP, Pitié-Salpêtrière Hospital, Paris, France), Maria Martinez (INSERM UMR 1220; and Paul Sabatier University, Toulouse, France),

Germany: Claudia Schulte (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, and DZNE, German Center for Neurodegenerative Diseases, Tübingen, Germany), Kathrin Brockmann (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, and DZNE, German Center for Neurodegenerative Diseases, Tübingen, Germany), Javier Simón-Sánchez (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, and DZNE, German Center for Neurodegenerative Diseases, Tübingen, Germany), Peter Heutink (DZNE, German Center for Neurodegenerative Diseases and Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany), Patrizia Rizzu (DZNE, German Center for Neurodegenerative Diseases), Manu Sharma (Centre for Genetic Epidemiology, Institute for Clinical Epidemiology and Applied Biometry, University of Tubingen, Germany), Thomas Gasser (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, and DZNE, German Center for Neurodegenerative

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Diseases, Tübingen, Germany), Susanne A. Schneider (Department of Neurology, Ludwig-Maximilians-University Munich, München, Germany)

United States of America: Mark R Cookson (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, USA), Sara Bandres-Ciga (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA), Cornelis Blauwendraat (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA), David W. Craig (Department of Translational Genomics, Keck School of Medicine, University of Southern California, Los Angeles, USA), Kimberley Billingsley (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA), Mary B. Makarious (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA), Derek P. Narendra (Inherited Movement Disorders Unit, National pInstitute of Neurological Disorders and Stroke, Bethesda, MD, USA), Faraz Faghri (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, USA; Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL, USA), J Raphael Gibbs (Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA), Dena G. Hernandez (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA), Kendall Van Keuren-Jensen (Neurogenomics Division, TGen, Phoenix, AZ USA), Joshua M. Shulman (Departments of Neurology, Neuroscience, and Molecular & Human Genetics, Baylor College of Medicine, Houston, Texas, USA; Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, Texas, USA), Hirotaka Iwaki (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA), Hampton L. Leonard (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA), Mike A. Nalls (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, USA; CEO/Consultant Data Tecnica International, Glen Echo, MD, USA), Laurie Robak

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

(Baylor College of Medicine, Houston, Texas, USA), Jose Bras (Center for Neurodegenerative Science, Van Andel Research Institute, Grand Rapids, Michigan, USA), Rita Guerreiro (Center for Neurodegenerative Science, Van Andel Research Institute, Grand Rapids, Michigan, USA), Steven Lubbe (Ken and Ruth Davee Department of Neurology and Simpson Querrey Center for Neurogenetics, Northwestern University Feinberg School of Medicine, Chicago, IL, USA), Steven Finkbeiner (Departments of Neurology and Physiology, University of California, San Francisco; Gladstone Institute of Neurological Disease; Taube/Koret Center for Neurodegenerative Disease Research, San Francisco, CA, USA), Niccolo E. Mencacci (Northwestern University Feinberg School of Medicine, Chicago, IL, USA), Codrin Lungu (National Institutes of Health Division of Clinical Research, NINDS, National Institutes of Health, Bethesda, MD, USA), Andrew B Singleton (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA), Sonja W. Scholz (Neurodegenerative Diseases Research Unit, National Institute of Neurological Disorders and Stroke, Bethesda, MD, USA), Xylena Reed (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA). Roy N. Alcalay (Department of Neurology, College of Physicians and Surgeons, Columbia University Medical Center, New York, NY, USA, Taub Institute for Research on Alzheimer's Disease and the Aging Brain, College of Physicians and Surgeons, Columbia University Medical Center, New York, NY, USA). Zbigniew K. Wszolek (Department of Neurology, Mayo Clinic Jacksonville, FL, USA), Ryan J. Uitti (Department of Neurology, Mayo Clinic Jacksonville, FL, USA), Owen A. Ross (Departments of Neuroscience & Clinical Genomics, Mayo Clinic Jacksonville, FL, USA), Francis P. Grenn (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA), Anni Moore (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA).

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Canada: Ziv Gan-Or (Montreal Neurological Institute and Hospital, Department of Neurology & Neurosurgery, Department of Human Genetics, McGill University, Montréal, QC, H3A 0G4, Canada), Guy A. Rouleau (Montreal Neurological Institute and Hospital, Department of Neurology & Neurosurgery, Department of Human Genetics, McGill University, Montréal, QC, H3A 0G4, Canada), Lynne Krohn (Montreal Neurological Institute and Hospital, Department of Neurology & Neurosurgery, Department of Human Genetics, McGill University, Montréal, QC, H3A 0G4, Canada), Kheireddin Mufti (Montreal Neurological Institute and Hospital, Department of Neurology & Neurosurgery, Department of Human Genetics, McGill University, Montréal, QC, H3A 0G4, Canada),

The Netherlands: Jacobus J van Hilten (Department of Neurology, Leiden University Medical Center, Leiden, Netherlands), Johan Marinus (Department of Neurology, Leiden University Medical Center, Leiden, Netherlands)

Spain: Astrid D. Adarmes-Gómez (Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville), Miquel Aguilar (Fundació Docència i Recerca Mútua de Terrassa and Movement Disorders Unit, Department of Neurology, University Hospital Mutua de Terrassa, Terrassa, Barcelona.), Ignacio Alvarez (Fundació Docència i Recerca Mútua de Terrassa and Movement Disorders Unit, Department of Neurology, University Hospital Mutua de Terrassa, Terrassa, Barcelona.),Victoria Alvarez (Hospital Universitario Central de Asturias, Oviedo), Francisco Javier Barrero (Hospital Universitario San Cecilio de Granada, Universidad de Granada), Jesús Alberto Bergareche Yarza (Instituto de Investigación Sanitaria Biodonostia, San Sebastián), Inmaculada Bernal-Bernal (Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville), Marta Blazquez (Hospital Universitario Central

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

de Asturias, Oviedo), Marta Bonilla-Toribio (Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville), Juan A. Botía (Universidad de Murcia, Murcia), María Teresa Boungiorno (Fundació Docència i Recerca Mútua de Terrassa and Movement Disorders Unit, Department of Neurology, University Hospital Mutua de Terrassa, Terrassa, Barcelona.) Dolores Buiza-Rueda (Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville), Ana Cámara (Hospital Clinic de Barcelona), Fátima Carrillo (Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville), Mario Carrión-Claro (Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville), Debora Cerdan (Hospital General de Segovia, Segovia), Jordi Clarimón (Memory Unit, Department of Neurology, IIB Sant Pau, Hospital de la Santa Creu i Sant Pau, Universitat Autònoma de Barcelona and Centro de Investigación Biomédica en Red en Enfermedades Neurodegenerativas (CIBERNED), Madrid),Yaroslau Compta (Hospital Clinic de Barcelona), Monica Diez-Fairen (Fundació Docència i Recerca Mútua de Terrassa and Movement Disorders Unit, Department of Neurology, University Hospital Mutua de Terrassa, Terrassa, Barcelona.), Oriol Dols-Icardo (Memory Unit, Department of Neurology, IIB Sant Pau, Hospital de la Santa Creu i Sant Pau, Universitat Autònoma de Barcelona, Barcelona, and Centro de Investigación Biomédica en Red en Enfermedades Neurodegenerativas (CIBERNED), Madrid), Jacinto Duarte (Hospital General de Segovia, Segovia), Raquel Duran (Centro de Investigacion Biomedica, Universidad de Granada, Granada), Francisco Escamilla-Sevilla (Hospital Universitario Virgen de las Nieves, Instituto de Investigación Biosanitaria de Granada, Granada), Mario Ezquerra (Hospital Clinic de Barcelona), Cici Feliz (Departmento de Neurologia, Instituto de Investigación Sanitaria Fundación Jiménez Díaz, Madrid, Spain), Manel Fernández (Hospital Clinic de Barcelona),

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Rubén Fernández-Santiago (Hospital Clinic de Barcelona), Ciara Garcia (Hospital Universitario Central de Asturias, Oviedo), Pedro García-Ruiz (Instituto de Investigación Sanitaria Fundación Jiménez Díaz, Madrid), Pilar Gómez-Garre (Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville), Maria Jose Gomez Heredia (Hospital Universitario Virgen de la Victoria, Malaga), Isabel Gonzalez-Aramburu (Hospital Universitario Marqués de Valdecilla-IDIVAL, Santander) ,Ana Gorostidi Pagola (Instituto de Investigación Sanitaria Biodonostia, San Sebastián), Janet Hoenicka (Institut de Recerca Sant Joan de Déu, Barcelona), Jon Infante (Hospital Universitario Marqués de Valdecilla-IDIVAL and University of Cantabria, Santander, and Centro de Investigación Biomédica en Red en Enfermedades Neurodegenerativas (CIBERNED), Silvia Jesús (Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville), Adriano Jimenez-Escrig (Hospital Universitario Ramón y Cajal, Madrid), Jaime Kulisevsky (Movement Disorders Unit, Department of Neurology, IIB Sant Pau, Hospital de la Santa Creu i Sant Pau, Universitat Autònoma de Barcelona, Barcelona, and Centro de Investigación Biomédica en Red en Enfermedades Neurodegenerativas (CIBERNED)), Miguel A. Labrador-Espinosa (Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville), Jose Luis Lopez-Sendon (Hospital Universitario Ramón y Cajal, Madrid), Adolfo López de Munain Arregui (Instituto de Investigación Sanitaria Biodonostia, San Sebastián), Daniel Macias (Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville), Irene Martínez Torres (Department of Neurology, Instituto de Investigación Sanitaria La Fe, Hospital Universitario y Politécnico La Fe, Valencia), Juan Marín (Movement Disorders Unit, Department of Neurology, IIB Sant Pau, Hospital de la Santa Creu i Sant Pau, Universitat Autònoma de Barcelona, Barcelona, and Centro de Investigación

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Biomédica en Red en Enfermedades Neurodegenerativas (CIBERNED)), Maria Jose Marti (Hospital Clinic Barcelona), Juan Carlos Martínez-Castrillo (Instituto Ramón y Cajal de Investigación Sanitaria, Hospital Universitario Ramón y Cajal, Madrid), Carlota Méndez-del-Barrio (Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville), Manuel Menéndez González (Hospital Universitario Central de Asturias, Oviedo), Marina Mata (Department of Neurology, Hospital Universitario Infanta Sofía, Madrid, Spain) Adolfo Mínguez (Hospital Universitario Virgen de las Nieves, Granada, Instituto de Investigación Biosanitaria de Granada), Pablo Mir (Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville), Elisabet Mondragon Rezola (Instituto de Investigación Sanitaria Biodonostia, San Sebastián), Esteban Muñoz (Hospital Clinic Barcelona), Javier Pagonabarraga (Movement Disorders Unit, Department of Neurology, IIB Sant Pau, Hospital de la Santa Creu i Sant Pau, Universitat Autònoma de Barcelona, Barcelona, and Centro de Investigación Biomédica en Red en Enfermedades Neurodegenerativas (CIBERNED)), Pau Pastor (Fundació Docència i Recerca Mútua de Terrassa and Movement Disorders Unit, Department of Neurology, University Hospital Mutua de Terrassa, Terrassa, Barcelona.), Francisco Perez Errazquin (Hospital Universitario Virgen de la Victoria, Malaga), Teresa Periñán-Tocino (Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville), Javier Ruiz-Martínez (Hospital Universitario Donostia, Instituto de Investigación Sanitaria Biodonostia, San Sebastián), Clara Ruz (Centro de Investigacion Biomedica, Universidad de Granada, Granada), Antonio Sanchez Rodriguez (Hospital Universitario Marqués de Valdecilla-IDIVAL, Santander), María Sierra (Hospital Universitario Marqués de Valdecilla-IDIVAL, Santander), Esther Suarez-Sanmartin (Hospital Universitario Central de Asturias, Oviedo), Cesar Tabernero (Hospital General de Segovia, Segovia), Juan Pablo Tartari

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

(Fundació Docència i Recerca Mútua de Terrassa and Movement Disorders Unit, Department of Neurology, University Hospital Mutua de Terrassa, Terrassa, Barcelona), Cristina Tejera-Parrado (Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville), Eduard Tolosa (Hospital Clinic Barcelona), Francesc Valldeoriola (Hospital Clinic Barcelona), Laura Vargas-González (Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville), Lydia Vela (Department of Neurology, Hospital Universitario Fundación Alcorcón, Madrid), Francisco Vives (Centro de Investigacion Biomedica, Universidad de Granada, Granada).

Austria: Alexander Zimprich (Department of Neurology, Medical University of Vienna, Austria)

Norway: Lasse Pihlstrom (Department of Neurology, Oslo University Hospital, Oslo, Norway), Mathias Toft (Department of Neurology and Institute of Clinical Medicine, Oslo University Hospital, Oslo, Norway)

Estonia: Pille Taba (Department of Neurology and Neurosurgery, University of Tartu, Tartu, Estonia)

Australia: Sulev Koks (Centre for Molecular Medicine and Innovative Therapeutics, Murdoch University, Murdoch, 6150, Perth, Western Australia; The Perron Institute for Neurological and Translational Science, Nedlands, 6009, Perth, Western Australia)

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Israel: Sharon Hassin-Baer (The Movement Disorders Institute, Department of Neurology and Sagol Neuroscience Center, Chaim Sheba Medical Center, Tel-Hashomer, 5262101, Ramat Gan, Israel, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel)

Finland: Kari Majamaa (Institute of Clinical Medicine, Department of Neurology, University of Oulu, Oulu, Finland; Department of Neurology and Medical Research Center, Oulu University Hospital, Oulu, Finland), Ari Siitonen (Institute of Clinical Medicine, Department of Neurology, University of Oulu, Oulu, Finland; Department of Neurology and Medical Research Center, Oulu University Hospital, Oulu, Finland), Pentti Tienari (Clinical Neurosciences, Neurology, University of Helsinki, Helsinki, Finland, Helsinki University Hospital, Helsinki, Finland)

Nigeria: Njideka U. Okubadejo (University of Lagos, Lagos State, Nigeria), Oluwadamilola O. Ojo (University of Lagos, Lagos State, Nigeria),

Kazakhstan: Coordinator - Rauan Kaiyrzhanov (Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Chingiz Shashkin (Kazakh National Medical University named after Asfendiyarov, Almaty, Kazakhstan), Nazira Zharkinbekova (South Kazakhstan Medical Academy, Shymkent, Kazakhstan), Vadim Akhmetzhanov (Astana Medical University, Astana Kazakhstan), Gulnaz Kaishybayeva (Scientific and practical center “Institute of neurology named after Smagul Kaishibayev”, Almaty, Kazakhstan), Altynay Karimova (Scientific and practical center “Institute of neurology named after Smagul Kaishibayev”, Almaty, Kazakhstan), Talgat Khaibullin (Semey Medical University, Semey, Kazakhstan).

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Ireland: Timothy L. Lynch (The Dublin Neurological Institute at the Mater Misericordiae University Hospital, Dublin, Ireland & School of Medicine and Medical Science, University College Dublin, Dublin, Ireland).

Conflicts of Interest: The authors have no potential conflicts of interest to report.

Data access: BMJ and AJN had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis

Sources of Funding: The Preventive Neurology Units is funded by the Barts Charity. Melanie Jensen received support from the Isaac Schapera Trust and Benjamin Jacobs is an NIHR Academic Clinical Fellow. The funding bodies had no role design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

References

1. de Rijk MC, Launer LJ, Berger K, et al. Prevalence of Parkinson's disease in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group. *Neurology.* 2000;54(11 Suppl 5):S21-23.
2. Cheng H-C, Ulane CM, Burke RE. Clinical progression in Parkinson disease and the neurobiology of axons. *Ann Neurol.* 2010;67(6):715-725.
3. Gruden MA, Sewell RDE, Yanamandra K, et al. Immunoprotection against toxic biomarkers is retained during Parkinson's disease progression. *Journal of neuroimmunology.* 2011;233(1):221-227.
4. Bas J, Calopa M, Mestre M, et al. Lymphocyte populations in Parkinson's disease and in rat models of parkinsonism. *J Neuroimmunol.* 2001;113(1):146-152.
5. Baba Y, Kuroiwa A, Uitti RJ, Wszolek ZK, Yamada T. Alterations of T-lymphocyte populations in Parkinson disease. *Parkinsonism Relat Disord.* 2005;11(8):493-498.
6. Stevens CH, Rowe D, Morel-Kopp M-C, et al. Reduced T helper and B lymphocytes in Parkinson's disease. *Journal of neuroimmunology.* 2012;252(1):95-99.
7. Niwa F, Kuriyama N, Nakagawa M, Imanishi J. Effects of peripheral lymphocyte subpopulations and the clinical correlation with Parkinson's disease. *Geriatrics & Gerontology International.* 2011;12(1):102-107.
8. Akil E, Bulut A, Kaplan I, Ozdemir HH, Arslan D, Aluclu MU. The increase of carcinoembryonic antigen (CEA), high-sensitivity C-reactive protein, and neutrophil/lymphocyte ratio in Parkinson's disease. *Neurol Sci.* 2015;36(3):423-428.
9. Tan E-K, Chao Y-X, West A, Chan L-L, Poewe W, Jankovic J. Parkinson disease and the immune system — associations, mechanisms and therapeutics. *Nature Reviews Neurology.* 2020.
10. Saiki M, Baker A, Williams-Gray CH, et al. Association of the human leucocyte antigen region with susceptibility to Parkinson's disease. *Journal of neurology, neurosurgery, and psychiatry.* 2010;81(8):890-891.
11. Nalls MA, Plagnol V, Hernandez DG, et al. Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet (London, England).* 2011;377(9766):641-649.
12. Bandres-Ciga S, Saez-Atienzar S, Kim JJ, et al. Large-scale pathway specific polygenic risk and transcriptomic community network analysis identifies novel functional pathways in Parkinson disease. *Acta Neuropathol.* 2020.
13. Sulzer D, Alcalay RN, Garretti F, et al. T cells from patients with Parkinson's disease recognize alpha-synuclein peptides. *Nature.* 2017;546(7660):656-661.
14. Bandres-Ciga S, Cookson MR. Alpha-synuclein triggers T-cell response. Is Parkinson's disease an autoimmune disorder? *Mov Disord.* 2017;32(9):1327.
15. Hui KY, Fernandez-Hernandez H, Hu J, et al. Functional variants in the LRRK2 gene confer shared effects on risk for Crohn's disease and Parkinson's disease. *Science translational medicine.* 2018;10(423).
16. Noyce AJ, Bestwick JP, Silveira-Moriyama L, et al. Meta-analysis of early nonmotor features and risk factors for Parkinson disease. *Ann Neurol.* 2012;72(6):893-901.
17. San Luciano M, Tanner CM, Meng C, et al. Nonsteroidal Anti-Inflammatory Use and LRRK2 Parkinson's Disease Penetrance. *Movement Disorders.* 2020;n/a(n/a).
18. Williams-Gray CH, Wijeyekoon R, Yarnall AJ, et al. Serum immune markers and disease progression in an incident Parkinson's disease cohort (ICICLE-PD). *Movement disorders : official journal of the Movement Disorder Society.* 2016;31(7):995-1003.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

19. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562(7726):203-209.
20. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet.* 2014;23(R1):R89-98.
21. Bandres-Ciga S, Noyce AJ, Traynor BJ. Mendelian Randomization-A Journey From Obscurity to Center Stage With a Few Potholes Along the Way. *JAMA Neurol.* 2019.
22. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol.* 2015;44(2):512-525.
23. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet.* 2017;13(11):e1007081.
24. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife.* 2018;7:e34408.
25. Vuckovic D, Bao EL, Akbari P, et al. The Polygenic and Monogenic Basis of Blood Traits and Diseases. *Cell.* 2020;182(5):1214-1231.e1211.
26. Hemani G, Tilling K, Davey Smith G. Correction: Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLOS Genetics.* 2017;13(12):e1007149.
27. Nalls MA, Blauwendraat C, Vallerga CL, et al. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet Neurol.* 2019;18(12):1091-1102.
28. Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *Am J Epidemiol.* 2013;178(7):1177-1184.
29. Yazdani S, Mariosa D, Hammar N, et al. Peripheral immune biomarkers and neurodegenerative diseases: A prospective cohort study with 20 years of follow-up. *Ann Neurol.* 2019;86(6):913-926.
30. Charlett A, Dobbs RJ, Dobbs SM, et al. Blood profile holds clues to role of infection in a premonitory state for idiopathic parkinsonism and of gastrointestinal infection in established disease. *Gut Pathog.* 2009;1(1):20.
31. Stevens CH, Rowe D, Morel-Kopp MC, et al. Reduced T helper and B lymphocytes in Parkinson's disease. *J Neuroimmunol.* 2012;252(1-2):95-99.
32. Umehara T, Oka H, Nakahara A, Matsuno H, Murakami H. Differential leukocyte count is associated with clinical phenotype in Parkinson's disease. *Journal of the Neurological Sciences.* 2020;409:116638.
33. Sommer A, Winner B, Prots I. The Trojan horse - neuroinflammatory impact of T cells in neurodegenerative diseases. *Molecular Neurodegeneration.* 2017;12(1):78.
34. Valiathan R, Ashman M, Asthana D. Effects of Ageing on the Immune System: Infants to Elderly. *Scandinavian Journal of Immunology.* 2016;83(4):255-266.
35. Sansoni P, Cossarizza A, Brianti V, et al. Lymphocyte subsets and natural killer cell activity in healthy old people and centenarians. *Blood.* 1993;82(9):2767-2773.
36. Costantini E, D'Angelo C, Reale M. The Role of Immunosenescence in Neurodegenerative Diseases. *Mediators of Inflammation.* 2018;2018:6039171.
37. Lin W-C, Tsai N-W, Huang Y-C, et al. Peripheral leukocyte apoptosis in patients with Parkinsonism: correlation with clinical characteristics and neuroimaging findings. *Biomed Res Int.* 2014;2014:635923-635923.
38. McKenna F, McLaughlin PJ, Lewis BJ, et al. Dopamine receptor expression on human T- and B-lymphocytes, monocytes, neutrophils, eosinophils and NK cells: a flow cytometric study. *Journal of neuroimmunology.* 2002;132(1):34-40.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

39. Besser MJ, Ganor Y, Levite M. Dopamine by itself activates either D2, D3 or D1/D5 dopaminergic receptors in normal human T-cells and triggers the selective secretion of either IL-10, TNFα or both. *Journal of neuroimmunology*. 2005;169(1):161-171.
40. Carr L, Tucker A, Fernandez-Botran R. In vivo administration of L-dopa or dopamine decreases the number of splenic IFN gamma-producing cells. *Journal of neuroimmunology*. 2003;137(1-2):87-93.
41. Mahlknecht P, Seppi K, Poewe W. The Concept of Prodromal Parkinson's Disease. *J Parkinsons Dis*. 2015;5(4):681-697.
42. Adams-Carr KL, Bestwick JP, Shribman S, Lees A, Schrag A, Noyce AJ. Constipation preceding Parkinson's disease: a systematic review and meta-analysis. *J Neurol Neurosurg Psychiatry*. 2016;87(7):710-716.
43. Heinzel S, Aho VTE, Suenkel U, et al. Gut microbiome signatures of risk and prodromal markers of Parkinson's disease. *bioRxiv*. 2019:2019.2012.2011.872481.
44. Eri T, Kawahata K, Kanzaki T, et al. Intestinal microbiota link lymphopenia to murine autoimmunity via PD-1+CXCR5−/dim B-helper T cell induction. *Scientific Reports*. 2017;7(1):46037.
45. Qiu X, Xiao Y, Wu J, Gan L, Huang Y, Wang J. C-Reactive Protein and Risk of Parkinson's Disease: A Systematic Review and Meta-Analysis. *Front Neurol*. 2019;10:384-384.
46. Hall S, Janelidze S, Surova Y, Widner H, Zetterberg H, Hansson O. Cerebrospinal fluid concentrations of inflammatory markers in Parkinson's disease and atypical parkinsonian disorders. *Sci Rep*. 2018;8(1):13276.
47. Lindqvist D, Hall S, Surova Y, et al. Cerebrospinal fluid inflammatory markers in Parkinson's disease--associations with depression, fatigue, and cognitive impairment. *Brain Behav Immun*. 2013;33:183-189.
48. Batty GD, Gale CR, Kivimaki M, Deary IJ, Bell S. Comparison of risk factor associations in UK Biobank against representative, general population based studies with conventional response rates: prospective cohort study and individual participant meta-analysis. *BMJ*. 2020;368:m131.
49. Britanova OV, Putintseva EV, Shugay M, et al. Age-Related Decrease in TCR Repertoire Diversity Measured with Deep and Normalized Sequence Profiling. *The Journal of Immunology*. 2014;192(6):2689.
50. Zidar DA, Al-Kindi SG, Liu Y, et al. Association of Lymphopenia With Risk of Mortality Among Adults in the US General Population. *JAMA Network Open*. 2019;2(12):e1916526-e1916526.
51. Noyce AJ, Bestwick JP, Silveira-Moriyama L, et al. PREDICT-PD: identifying risk of Parkinson's disease in the community: methods and baseline results. *J Neurol Neurosurg Psychiatry*. 2014;85(1):31-37.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Figure legends

Figure 1: Association of blood cell marker and inflammatory markers with incident PD in UKB. Betas and 95% CIs are shown from multivariable logistic regression models of the form PD ~ Age + Sex + Ethnicity + Townsend deprivation score + blood biomarker. Estimates shown here are for Z-score standardised variables (x-mean / sd) to allow comparison between effect sizes.

Figure 2: Mendelian randomisation analysis of the effect of lymphocyte count on PD risk. A: scatter plot showing SNP associations with lymphocyte count from our GWAS (x) and with the per-allele log(OR) for PD (estimates from the largest PD GWAS published, Meta5, excluding UKB participants). The model fit lines indicate MR estimates (of the 'causal effect) from different MR methods. B: MR estimates from various methods using the 'Mixture of Experts approach'. The y axis shows different MR methods and different approaches for filtering SNPs to be included in the genetic instrument (heterogeneity filtering, HF; directionality filtering, DF). Estimates are coloured and ordered by the 'MOE' statistic, which is similar to an area under the curve statistic in that it quantifies that ability of a given MR method to distinguish a true effect from the absence of a true effect. MOE statistics closer to 1 indicate a higher likelihood that the given MR method will give an accurate estimate for the given dataset.

Figure 1

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

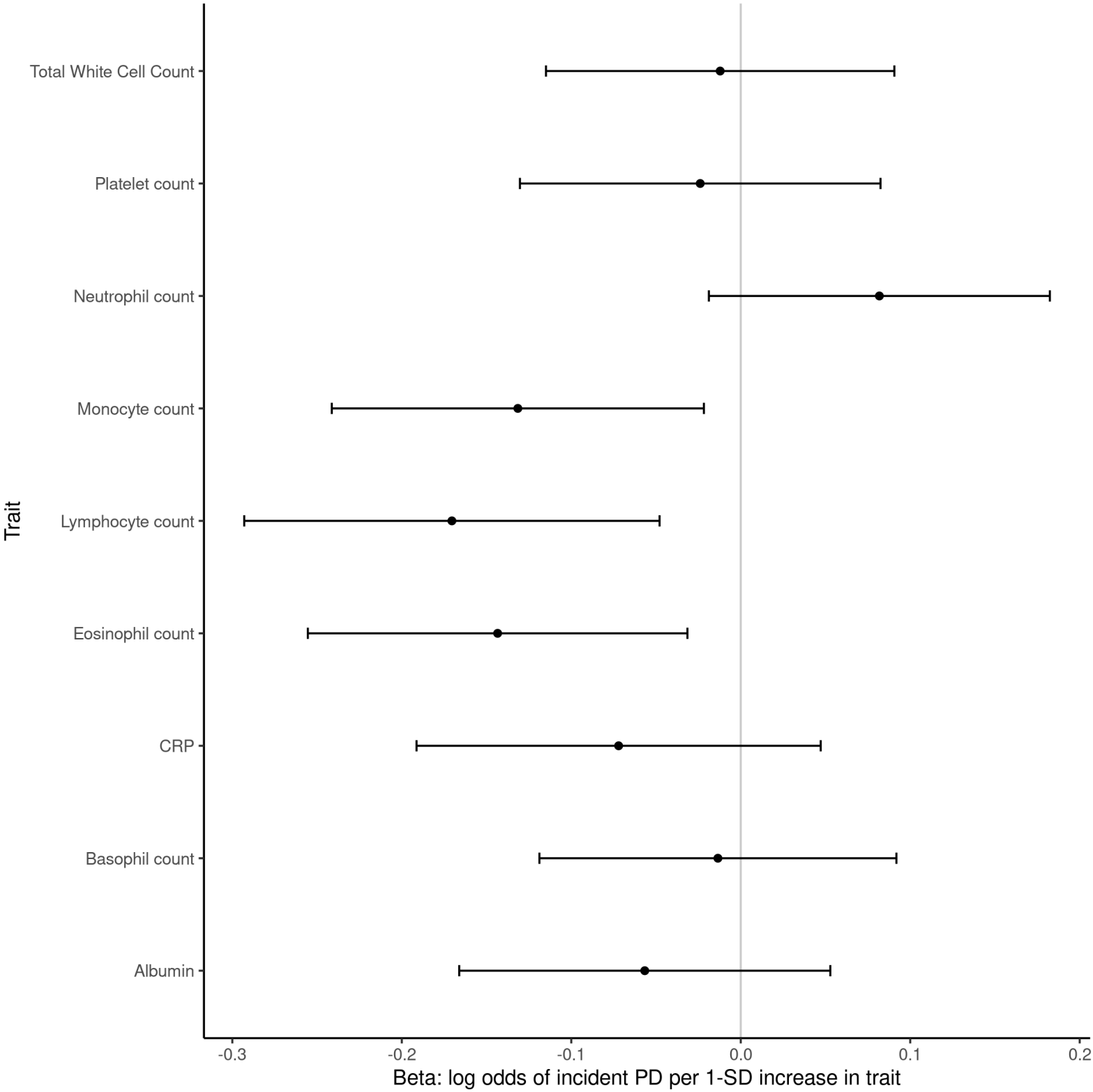


Figure 2A

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

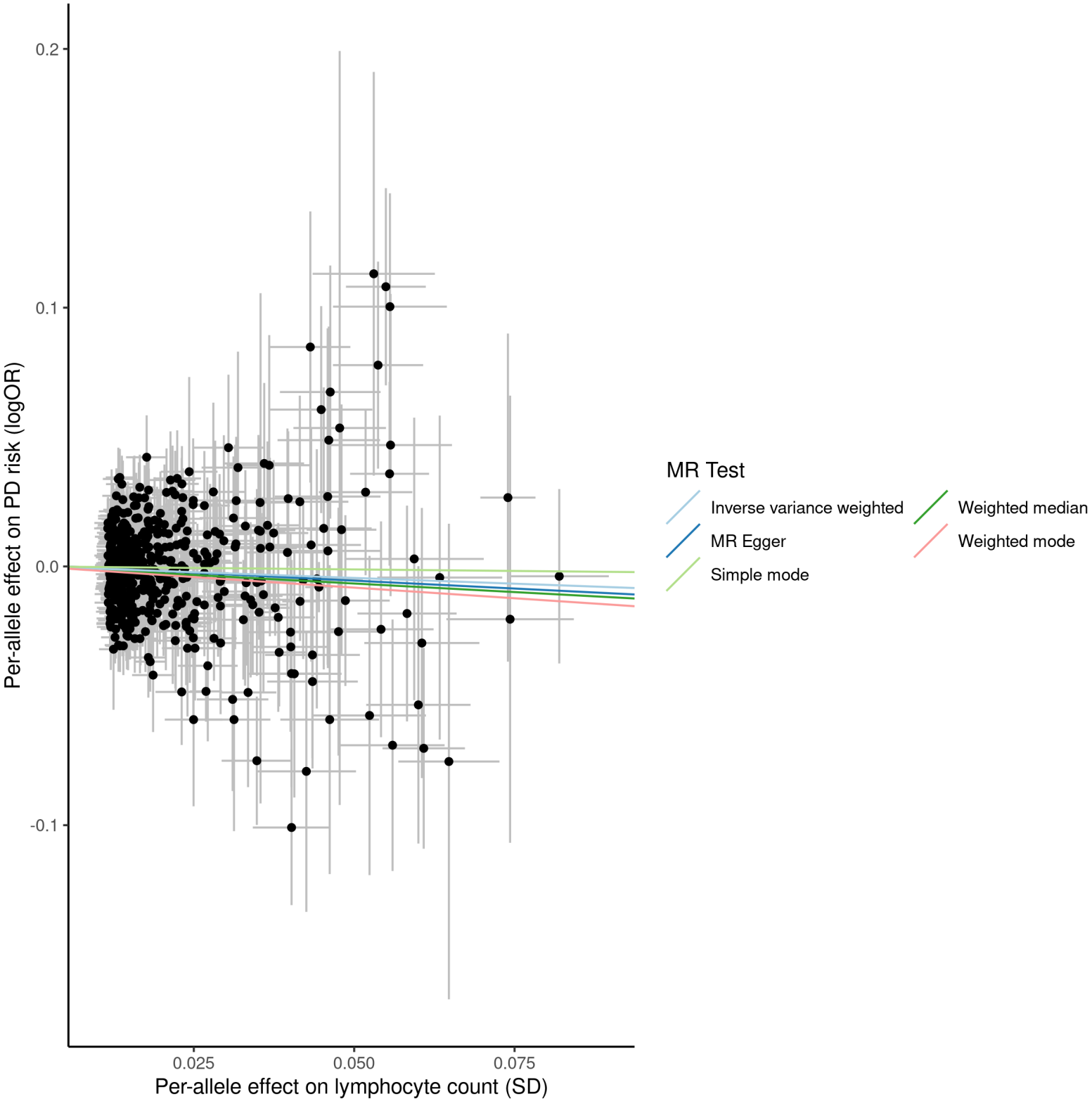
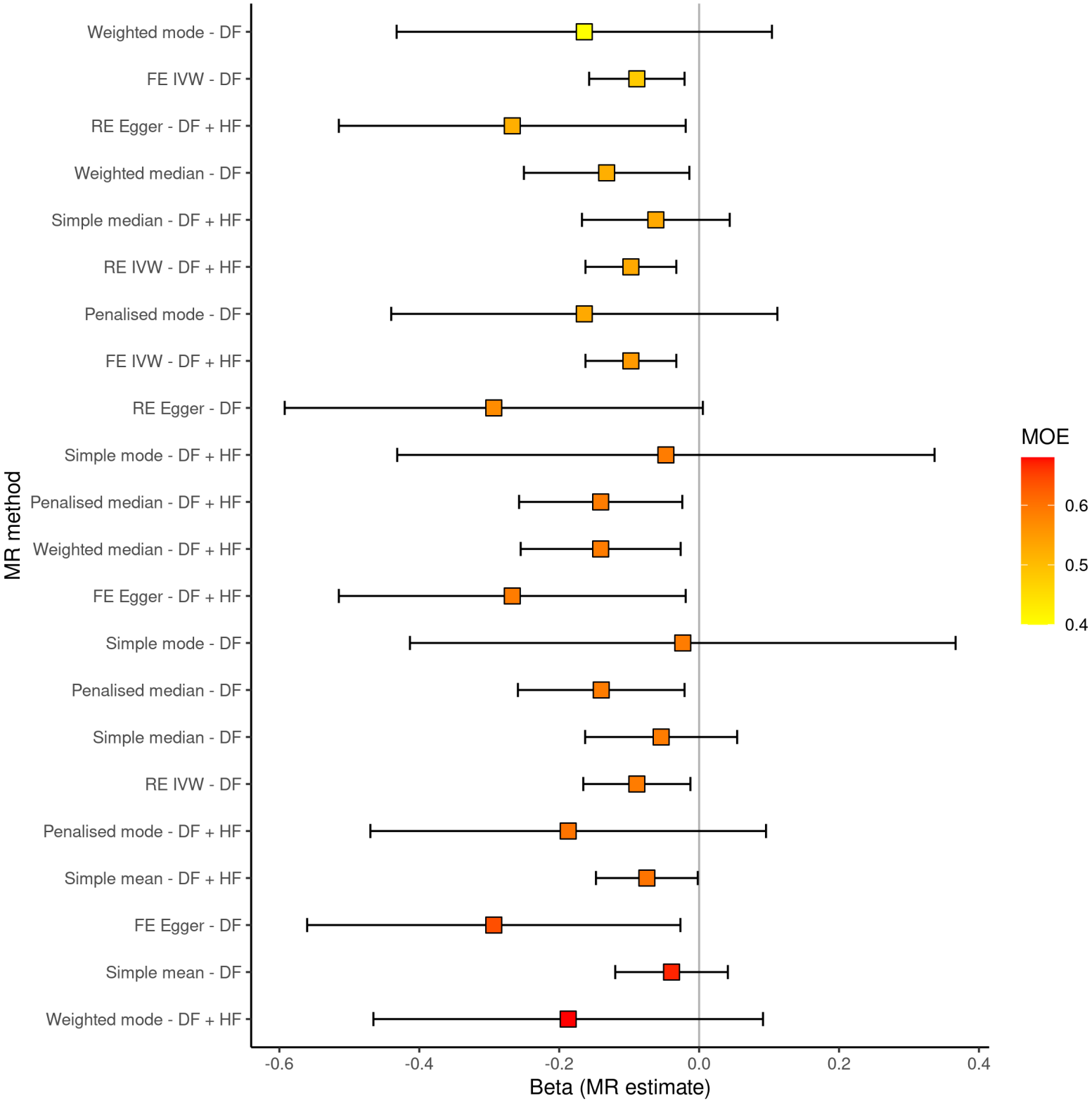


Figure 2B

medRxiv preprint doi: https://doi.org/10.1101/2020.09.13.20189530; this version posted September 24, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .