

Towards a phenome-wide view of Parkinson's disease

Authors

Abby L. Olsen, MD PhD^{1,2,3}, Joseph J. Locascio, PhD^{1,2,4}, Idil Tuncali, MS^{1,2,3}, Nada Laroussi^{1,2,3}, Elena Abatzis^{1,2,3}, Polina Kamenskaya^{1,2,3}, Yuliya Kuras, PhD^{1,2,3}, Tom Yi^{1,2,3}, Aleksandar Videnovic, MD, MSc⁴, Michael T. Hayes, MD³, Gary P. H. Ho, MD, PhD³, Jordan Paulson, MD³, Vikram Khurana, MD, PhD³, Todd M. Herrington, MD, PhD⁴, Bradley T. Hyman, MD, PhD⁴, Dennis J. Selkoe, MD³, John H. Growdon, MD⁴, Stephen N. Gomperts, MD, PhD⁴, Trond Riise⁵, PhD, Michael A. Schwarzschild, MD, PhD⁴, Albert Y. Hung, MD, PhD^{3,4}, Anne-Marie Wills, MD, MPH⁴, and Clemens R. Scherzer, MD^{1,2,3,4,6*}

1 APDA Center for Advanced Parkinson Research, Harvard Medical School, Brigham and Women's Hospital, Boston, MA 02115, USA

2 Precision Neurology Program, Harvard Medical School, Brigham and Women's Hospital, Boston, MA 02115, USA

3 Department of Neurology, Brigham and Women's Hospital, Boston, MA 02115, USA

4 Department of Neurology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA

5 Department of Global Public health and Primary Care, University of Bergen, Norway.

6 Aligning Science Across Parkinson's (ASAP) Collaborative Research Network, Chevy Chase, MD, 20815

Background and Objectives: Many studies have examined the relation between PD and environmental variables serially --- one candidate association at a time. In the real world however, both environmental exposures and patients are much more complex, including correlated environmental exposures, polypharmacy, and complex comorbidities. Here we begin to characterize a holistic view of environmental, health, and pharmacological traits linked to patients with PD.

Methods: The Harvard Biomarkers Study (HBS) is a large case-control study of PD patients and healthy controls that includes an extensive questionnaire covering past medical and social history data and is thus well-suited for such an exploratory study. Sixty-four environmental, pharmacological, and clinical features were evaluated for associations with PD using logistic regression analysis with backward elimination.

Results: Male gender, coronary artery disease, depression, anxiety, restless leg syndrome, head trauma, ibuprofen use, co-enzyme Q10 use, and vitamin D supplementation were significantly positively associated with PD. By contrast, asthma/chronic obstructive pulmonary disease (COPD), naproxen, ezetimibe, and smoking were significantly negatively associated with PD.

Discussion: This study shows that unbiased, data-rich exploration of the Parkinson phenome has the promise to uncover, prioritize, and clarify associations between environment, multi-system health phenotypes, and PD in a patient-centric manner. Associations with coronary artery

disease, mood disorders, and the cholesterol-absorption inhibitor ezetimibe were revealed that have been largely neglected in traditional hypothesis-driven investigations. Interestingly, asthma/COPD was inversely associated with PD, and this was independent of smoking history. Furthermore, well-established associations were confirmed for male gender, smoking, head trauma, and restless legs syndrome.

Abbreviations

PD = Parkinson's disease, HBS = Harvard Biomarkers Study

Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disorder. Although a minority of cases are familial, the underlying disease driver for most so-called idiopathic PD cases is unknown. PD is likely to arise through a complex interplay of genetic¹⁻³ and environmental factors^{4,5}. Many traditional epidemiological studies have examined the relation between PD, health data, and environmental variables serially⁶. In the real world, however, both environmental exposures and patients are much more complex, including phenotypic diversity, complex comorbidities, and polypharmacy. Here we begin to characterize the phenome of PD, that is, a holistic picture of environmental, clinical, and pharmacological traits linked to patients with PD.

We developed the Harvard Biomarkers Study (HBS) as a resource for identifying genes, and biomarkers for PD^{1,7-9}. HBS includes an extensive questionnaire regarding past medical history, medication and supplement use history, social history, and environmental exposures. This includes data on exposure to some previously reported putative risk or protective factors (e.g. smoking, pesticides) (Figure 1), though not all (e.g. dairy intake, exercise).

In this report, we perform logistic regression with backward elimination to determine which of these health variables are positively or negatively associated with PD. With a few notable exceptions^{10,11}, most epidemiologic studies have largely examined one or a few variables at a time. Without adjusting predictor variables for each other in a multivariate model, complex inter-correlations can produce spurious associations, suppression effects, and other obscuring statistical phenomena¹². Thus, our study represents a rare attempt to more comprehensively characterize the phenome of PD.

Methods

Harvard Biomarkers Study

The Harvard Biomarkers Study (HBS) is a case-control study including 3,000 patients with various neurodegenerative diseases as well as healthy controls (HC). Informed consent was obtained for all participants. The study protocol was approved by the institutional review board of Mass General Brigham. The HBS questionnaire is included as Supplemental Table 1. For more information on the HBS see: www.bwhparkinsoncenter/biobank

This analysis was limited to a subgroup of the larger HBS, consisting of 1224 total subjects: 933 with PD and 291 healthy controls. Healthy controls consisted chiefly of spouses, friends, and non-blood relatives, who accompanied PD patients to office visits. Diagnosis of PD was made by a board-certified neurologist with fellowship training in movement disorders. Subjects completed a detailed questionnaire including information on past medical history, current and prior medication use, nutritional supplement use, environmental exposures, Parkinson's disease risk factors, and social history (Figure 1). Each item in the questionnaire is phrased as a binary YES/NO question, which is followed up by quantitative questions in some instances. For example, the question "Do you drink caffeinated coffee?" is followed by questions asking how many cups per day on average and whether the consumption has changed over the past 10 years. For purposes of this initial study, we have limited the analysis to the single binary YES/NO question for each variable. Additionally, we have examined data from a single

timepoint, that of enrollment in the study. At the time of enrollment, average disease duration was 3.8 years (Table 1).

Statistical differences in demographic data (age, sex, and race) between the cases and controls were determined using a Satterwaithe t-test, Chi-square test, or Fisher's exact test, as appropriate (Table 1). Included in the table are the mean Unified Parkinson's Disease Rating Scale (UPDRS) and the Mini-Mental State Exam (MMSE) scores. The UPDRS is a 4-part clinical rating scale used to measure severity and progression of PD. In the UPDRS motor subscale (part 3), motor signs were assessed by a trained examiner. The MMSE is a 30-question cognitive battery featuring questions on orientation, registration, attention and calculation, recall, language, and copying.

Variable correlation

To get a preliminary sense of which subsets of the predictor variables are statistically similar and distinct in this data set, we subjected the Pearson correlation matrix of the entire set of 64 variables to a hierarchical clustering algorithm (employing the Corrplot package in R, version 1.1.423 and the graphical software GraphPad Prism, version 8.43).

Logistic regression

Multivariate logistic regression with backward elimination was performed in SAS version 9.4, where cases versus controls (PD or healthy controls) served as the binary dependent variable and the initial predictor set included the 64 binary phenome variables, in addition to the covariates of sex and age. Each predictor variable is statistically adjusted for all other variables still in the model at each given step in the backward elimination. The threshold for retaining predictor variables was $p < 0.05$. After excluding subjects with missing values, environmental data from 481 PD and 139 healthy controls (620 total subjects) were included in the final logistic regression model (Table 2, Figure 2).

For the supplemental analyses, we determined the univariate p-value for the relation of each individual predictor variable to group status (PD versus healthy controls), (Column 2, Supplemental Table 2). We also determined the p-value for each individual variable controlled for sex and age (but not all other environmental variables) (Column 3, Supplemental Table 2) using logistic regression. Finally, we include the p-value for each individual variable, controlling for sex and age as well as multiple testing using the Sidak step-down method or False Discovery Rate method (Columns 4-5, Supplemental Table 2).

Results

Demographics of the HBS cohort are shown in Table 1. The goal of this study was to evaluate the feasibility and promise of comprehensively characterizing the "PD phenome" by examining 64 variables spanning the categories of past medical history, medications and supplements, exposures previously reported to be PD risk factors, and social history factors (Figure 1).

Because some variables may be correlated with others, we first explored potential correlations using a correlation matrix (Figure 2A). Overall correlation between variables was low: the largest positive correlation coefficient was 0.43 between calcium and vitamin D

supplement use, and the largest negative correlation coefficient was -0.16 between full dose aspirin (325 mg) and baby aspirin (81 mg) use. However, the hierarchical clustering analysis does demonstrate that several sets of variables are correlated with P values below 0.01, forming clusters of pharmacological, disease, and exposure features that tag biologically related phenomena. The largest clusters in our dataset include a linked sleep and mood disorder cluster, a metabolic syndrome cluster, and a vitamin use cluster (Figure 2B).

We performed a logistic regression analysis with backward elimination to determine which variables were associated with PD. In total, the input included 64 binary health variables (Figure 1) as well as age and sex, assessed in 291 healthy controls and 933 subjects with PD. After the backward elimination and excluding subjects with missing values, 481 PD and 139 HC were included in the final model. The analysis identified 14 statistically significant variables (each significant after simultaneously adjusting for the other 13) after backward elimination (Table 2, Figure 3). Male sex, coronary artery disease (CAD), depression, anxiety, restless leg syndrome, head trauma, ibuprofen use, co-enzyme Q 10 use, and vitamin D supplement use were positively associated with PD. Older age was also associated with PD; odds ratios for each 1-year, 5-year, or 10-year increment in age are shown in Figure 3. Asthma/chronic obstructive pulmonary disease (COPD), naproxen use, ezetimibe use, and smoking were inversely associated with PD. The prevalence of each variable in PD versus healthy controls is shown in Table 2. Additional statistical analyses for all variables are shown in Supplemental Table 2, including the univariate P -value and the P -value after adjustment for sex and age only.

Among the environmental variables linked to PD, small but statistically significant correlations were found between restless leg syndrome and depression, restless leg syndrome and coronary artery disease, anxiety and depression, depression and asthma/COPD, ezetimibe use and coronary artery disease, and asthma/COPD and ibuprofen use (Figure 4).

Many patients with PD may have more than one factor that is associated with PD, such as being diagnosed with both restless leg syndrome and depression. Indeed, 38% of PD subjects (and 15% of healthy control subjects) were positive for two or more of the eight phenome variables that were identified as positively associated with PD in the logistic regression (Figure 5A). We examined the prevalence and odds ratio (OR) for each pairwise combination of variables positively associated with PD, in the PD population (Figure 5B-C). Among the pairwise combinations, the most prevalent variable found in combination with others was depression, and the single most prevalent combination was having both depression and anxiety (OR 3.00). Furthermore, depression became increasingly enriched in the PD subjects as their total number of positive variables grew. That is, the higher the number of positive variables present in a given subject, the more likely one of those variables was to be depression (Figure 5D). In total, depression was twice as prevalent in PD subjects compared to healthy controls (Table 2).

Discussion

This study reveals an initial multi-dimensional and data-rich view of phenome associations with PD. The goal of this study was to evaluate the feasibility and promise of this approach. As this is a case-control study, the results are not intended to imply direction of causality. That is, while some associations may arise because a phenotypic variable affects risk of PD, others likely arise because of having a diagnosis of PD affects risk of the phenotypic variable.

One strength of our analysis lies in replicating some previously reported associations, including the inverse association between smoking and PD and positive association between head trauma and PD. Of all environmental factors that have been inversely associated with PD, the relationship between smoking and PD is perhaps the best established. This was documented as early as the late 1960's¹³ and has been reproduced by numerous subsequent epidemiologic studies¹⁴. The mechanisms underlying this association are unknown, though there is ample evidence for a protective role of nicotine in dopaminergic neurons in animal models of PD¹⁵⁻¹⁸, but this has not been successfully confirmed in human clinical trials¹⁹⁻²¹. Interestingly, in our data, smoking was associated with a larger metabolic syndrome cluster, which also includes atorvastatin and pravastatin use (Fig. 2B). Several epidemiologic studies have examined statin use in PD^{10,22-25}, with mixed results. Thus, these clusters highlight the need for epidemiologic studies to consider related variables in linked groups.

The evidence for a positive association between head trauma and PD has been more mixed. While several studies including a 2013 meta-analysis have demonstrated an association²⁶⁻²⁸, multiple large population level Scandinavian studies have not²⁹⁻³¹. One potential explanation for the varied results are gene-environment interactions, as head trauma has been suggested to influence risk of PD specifically in subjects with higher genetic risk³²⁻³⁴. Another complicating factor is recall bias as well as the timing of the injury. For example, Rughjerg et al³⁵ found that any association between head trauma and PD could be explained by injuries in the months leading up to diagnosis, suggesting that head injury occurs in the setting of prodromal PD. In contrast, Taylor et al³⁶ excluded cases of trauma in the 10 years preceding the diagnosis of PD and found that head injury early in life was associated with PD. Thus, further studies are needed.

Although this analysis confirmed relationships between smoking and head trauma and PD, other previously reported associations, such as the inverse association between caffeine and PD, did not reach significance in the current analysis. Of note, in a prior study of the HBS cohort, we reported an inverse relationship between caffeine and PD⁹. The discrepancy between that study and the current analysis is explained by the fact that we here simplistically reduced all variables to binary yes/no exposures (considering the large number of variables analyzed), whereas in our prior analysis, caffeine intake was quantified. This highlights the need for future studies examining phenome variables holistically and quantitatively rather than categorically.

Beyond addressing these previously reported associations, our analysis demonstrates novel associations that warrant future study. First, we found an inverse relationship between PD and a past medical history of asthma or COPD. To our knowledge, this is the first report of such an association, though there is one study in a Taiwanese population demonstrating *increased* risk of PD in the setting of asthma or COPD^{37,38}.

It is important to consider whether smoking could explain the inverse association between asthma/COPD and PD. Interestingly, in our deeply phenotyped cohort, a diagnosis of asthma/COPD was not associated with smoking (P value = 0.51 with correlation coefficient of 0.019, Figure 2A); rather, smoking was associated with a larger metabolic syndrome cluster (Figure 2B). This inverse association between asthma/COPD and PD is interesting in light of prior work demonstrating that beta2-adrenoreceptor (beta2-AR) agonists, used to treat asthma and COPD, were associated with repression of alpha-synuclein expression and reduced risk of developing PD, whereas the beta 2 antagonist propranolol (when given for cardiovascular indication and excluding use for essential tremor) was associated with increased alpha-synuclein expression and increased risk of developing PD³⁹⁻⁴¹. Consistent with this, a recent meta-analysis of eight studies indicated that beta2-AR agonist use was associated with reduced PD risk (RR =

0.859, 95% confidence interval [CI] 0.741–0.995), while beta2-AR antagonist use was associated with an increased risk of PD (RR = 1.490, 95% CI, 1.195 to 1.857). In HBS, use of many oral medications was carefully recorded, but inhaled medications such as beta2-adrenoreceptor agonists used for asthma were not recorded. Thus, the current analysis cannot further clarify, whether asthma medications or asthma/COPD diagnosis mediates this observed association.

Another striking finding in our analysis was the strong prevalence of a history of depression in our PD population (Figure 5). Depression was especially enriched in PD subjects who had multiple positive variables associated with PD (as identified by our logistic regression analysis). The interplay between mood disorders and PD is complex, with some evidence suggesting that these represent either pre-motor⁴² or very early non-motor⁴³ co-comorbidities. These data highlight the importance of screening PD patients for mood disorders, which have a large impact on quality of life and are amenable to treatment⁴⁴.

Finally, an unexpected finding in our study pertains to non-steroidal anti-inflammatory drug (NSAID) use. Non-aspirin NSAIDs have been associated with reduced risk of PD in some^{10,45} but not all studies^{45,46}. In our cohort, naproxen use was negatively associated with PD, whereas ibuprofen use was positively associated with PD (Figure 3, Table 2). (We found no significant association between aspirin or celecoxib use and PD). One potential explanation for this apparent discrepancy is that ibuprofen use was most highly correlated with sertraline use (coefficient 0.14), which is in turn correlated with depression (0.18) and anxiety (0.21) (Figure 2), which were strongly associated with PD (Table 2). This example highlights the complicated relationship between different environmental exposures as well as their presumed mechanism of action: NSAIDs are not only anti-inflammatory but affect numerous cellular processes⁴⁷ and are highly effective analgesics, which is relevant given the high co-morbidity between pain and depression⁴⁸.

In summary, here we have provided an initial comprehensive characterization of the PD phenome using a cohort from the two major Harvard-affiliated hospitals. Our results confirm some previously reported associations as well as highlight other novel associations. Many of the health variables we have examined here are modifiable, meaning that these results may someday have implications for personalized medicine⁴⁹. Future work will require mechanistic studies to identify gene-environment interactions, to determine which factors are truly causative, and to discover whether modifying them has a neuroprotective or symptomatic benefit. As the only patient cohort with this extensive collection of environmental exposure data combined with whole genome sequencing, the Harvard Biomarkers Study represents an essential resource for undertaking these future studies.

Acknowledgements

The Harvard Biomarkers Study (HBS) (<https://www.bwhparkinsoncenter.org>) is a collaborative initiative of Brigham and Women's Hospital and Massachusetts General Hospital, co-directed by Dr. Clemens Scherzer and Dr. Bradley T. Hyman. The **HBS Study Investigators are:**

Co-Directors: Brigham and Women's Hospital: Clemens R. Scherzer, Massachusetts General Hospital: Bradley T. Hyman; Investigators and Study Coordinators: Brigham and Women's Hospital: Yuliya Kuras, Nada Laroussi, Elena Abatzias, Polina Kamenskya. Study Investigators: Brigham and Women's Hospital: Michael T. Hayes,

Aleksandar Videnovic, Nutan Sharma, Vikram Khurana, Claudio Melo De Gusmao, Reisa Sperling; Massachusetts General Hospital: John H. Growdon, Michael A. Schwarzschild, Albert Y. Hung, Alice W. Flaherty, Deborah Blacker, Anne-Marie Wills, Steven E. Arnold, Ann L. Hunt, Nichte I. Mejia, Anand Viswanathan, Stephen N. Gomperts, Mark W. Albers, Maria Allora-Palli, David Hsu, Alexandra Kimball, Scott McGinnis, John Becker, Randy Buckner, Thomas Byrne, Maura Copeland, Bradford Dickerson, Matthew Frosch, Theresa Gomez-Isla, Steven Greenberg, Julius Hedden, Elizabeth Hedley-Whyte, Keith Johnson, Raymond Kelleher, Aaron Koenig, Maria Marquis-Sayagues, Gad Marshall, Sergi Martinez-Ramirez, Donald McLaren, Olivia Okereke, Elena Ratti, Christopher William, Koene Van Dij, Shuko Takeda, Anat Stemmer-Rachaminov, Jessica Kloppenburg, Catherine Munro, Rachel Schmid, Sarah Wigman, Sara Wlodarczyk; Data Coordination: Brigham and Women's Hospital: Thomas Yi; Biobank Management Staff: Brigham and Women's Hospital: Idil Tuncali.”

We thank all study participants and their families for their invaluable contributions. HBS was seeded by generous support from the Harvard NeuroDiscovery Center, with partial contributions from APDA, the Michael J Fox Foundation, NINDS U01NS082157, U01NS100603, and the Massachusetts Alzheimer's Disease Research Center NIA P50AG005134.

C.R.S.'s work is supported by NIH grants U01NS095736, U01NS100603, R01AG057331, and R01NS115144, and the American Parkinson Disease Association Center for Advanced Parkinson Research.

The study is funded by the joint efforts of The Michael J. Fox Foundation for Parkinson's Research (MJFF) and the Aligning Science Across Parkinson's (ASAP) initiative. MJFF administers the grant [ASAP-000301] on behalf of ASAP and itself.

Author Contributions

ALO: manuscript writing, data analysis, interpretation of results. JLL, TR: statistical consultation, manuscript revision, interpretation of results. CRS: HBS design, project design, manuscript revision, interpretation of results. YK, PK, TY, EA, IT, NL: clinical data management and sample collection. AV, MTH, GPH, JP, VK, TMH, BH, DS, JHG, STG, MAS, AYH, A-M W: referring patients to the study, manuscript review.

Conflicts of Interest

CRS is named as co-inventor on a US patent application on sphingolipids biomarkers that is jointly held by Brigham & Women's Hospital and Sanofi. CRS has consulted for Sanofi Inc. and Calico; has collaborated with Pfizer, Opko, and Proteome Sciences, Genzyme Inc., and Lysosomal Therapies; is on the Scientific Advisory Board of the American Parkinson Disease Association; has served as Advisor to the Michael J. Fox Foundation, NIH, Department of Defense, and Google; has received funding from the NIH, the U.S. Department of Defense, the Michael J. Fox Foundation, and the American Parkinson Disease Association.

References

1. Liu G, Peng J, Liao Z, et al. Genome-wide survival study identifies a novel synaptic locus and polygenic score for cognitive progression in Parkinson's disease. *Nat Genet.* 2021;53(6):787-793. doi:10.1038/s41588-021-00847-6
2. 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. doi:10.1016/S1474-4422(19)30320-5
3. Zheng B, Liao Z, Locascio JJ, et al. PGC-1 α , a potential therapeutic target for early intervention in Parkinson's disease. *Sci Transl Med.* 2010;2(52):52ra73. doi:10.1126/scitranslmed.3001059
4. Kalia LV, Lang AE. Parkinson's disease. *Lancet.* 2015;386(9996):896-912. doi:10.1016/S0140-6736(14)61393-3
5. Ascherio A, Schwarzschild MA. The epidemiology of Parkinson's disease: risk factors and prevention. *Lancet Neurol.* 2016;15(12):1257-1272. doi:10.1016/S1474-4422(16)30230-7
6. Bellou V, Belbasis L, Tzoulaki I, Evangelou E, Ioannidis JPA. Environmental risk factors and Parkinson's disease: An umbrella review of meta-analyses. *Parkinsonism Relat Disord.* 2016;23:1-9. doi:10.1016/j.parkreldis.2015.12.008
7. Ding H, Dhima K, Lockhart KC, et al. Unrecognized vitamin D3 deficiency is common in Parkinson disease: Harvard Biomarker Study. *Neurology.* 2013;81(17):1531-1537. doi:10.1212/WNL.0b013e3182a95818
8. Locascio JJ, Eberly S, Liao Z, et al. Association between α -synuclein blood transcripts and early, neuroimaging-supported Parkinson's disease. *Brain.* 2015;138(Pt 9):2659-2671. doi:10.1093/brain/awv202
9. Bakshi R, Macklin EA, Hung AY, et al. Associations of Lower Caffeine Intake and Plasma Urate Levels with Idiopathic Parkinson's Disease in the Harvard Biomarkers Study. *J Parkinsons Dis.* 2020;10(2):505-510. doi:10.3233/JPD-191882
10. 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. doi:10.1002/ana.23687
11. Heilbron K, Noyce AJ, Fontanillas P, et al. The Parkinson's phenome-traits associated with Parkinson's disease in a broadly phenotyped cohort. *NPJ Parkinsons Dis.* 2019;5:4. doi:10.1038/s41531-019-0077-5

12. Locascio JJ, Lee J, Meltzer HY. Importance of adjusting for correlated concomitant variables in psychiatric research. *Psychiatry Res.* 1988;23(3):311-327. doi:10.1016/0165-1781(88)90022-4
13. Nefzger MD, Quadfasel FA, Karl VC. A retrospective study of smoking in Parkinson's disease. *Am J Epidemiol.* 1968;88(2):149-158. doi:10.1093/oxfordjournals.aje.a120874
14. Breckenridge CB, Berry C, Chang ET, Sielken RL, Mandel JS. Association between Parkinson's Disease and Cigarette Smoking, Rural Living, Well-Water Consumption, Farming and Pesticide Use: Systematic Review and Meta-Analysis. *PLoS One.* 2016;11(4):e0151841. doi:10.1371/journal.pone.0151841
15. Yang J, Lv DJ, Li LX, et al. Nicotine improved the olfactory impairment in MPTP-induced mouse model of Parkinson's disease. *Neurotoxicology.* 2019;73:175-182. doi:10.1016/j.neuro.2019.02.008
16. Kardani J, Sethi R, Roy I. Nicotine slows down oligomerisation of α -synuclein and ameliorates cytotoxicity in a yeast model of Parkinson's disease. *Biochim Biophys Acta Mol Basis Dis.* 2017;1863(6):1454-1463. doi:10.1016/j.bbadis.2017.02.002
17. Dunn AR, Stout KA, Ozawa M, et al. Synaptic vesicle glycoprotein 2C (SV2C) modulates dopamine release and is disrupted in Parkinson disease. *Proc Natl Acad Sci USA.* 2017;114(11):E2253-E2262. doi:10.1073/pnas.1616892114
18. Nicholatos JW, Francisco AB, Bender CA, et al. Nicotine promotes neuron survival and partially protects from Parkinson's disease by suppressing SIRT6. *Acta Neuropathol Commun.* 2018;6(1):120. doi:10.1186/s40478-018-0625-y
19. Villafane G, Thiriez C, Audureau E, et al. High-dose transdermal nicotine in Parkinson's disease patients: a randomized, open-label, blinded-endpoint evaluation phase 2 study. *Eur J Neurol.* 2018;25(1):120-127. doi:10.1111/ene.13474
20. Ebersbach G, Stöck M, Müller J, Wenning G, Wissel J, Poewe W. Worsening of motor performance in patients with Parkinson's disease following transdermal nicotine administration. *Mov Disord.* 1999;14(6):1011-1013. doi:10.1002/1531-8257(199911)14:6<1011::aid-mds1016>3.0.co;2-f
21. Vieregge A, Sieberer M, Jacobs H, Hagenah JM, Vieregge P. Transdermal nicotine in PD: a randomized, double-blind, placebo-controlled study. *Neurology.* 2001;57(6):1032-1035. doi:10.1212/wnl.57.6.1032
22. Jeong SM, Jang W, Shin DW. Association of statin use with Parkinson's disease: Dose-response relationship. *Mov Disord.* 2019;34(7):1014-1021. doi:10.1002/mds.27681

23. Rozani V, Giladi N, El-Ad B, et al. Statin adherence and the risk of Parkinson's disease: A population-based cohort study. *PLoS One*. 2017;12(4):e0175054. doi:10.1371/journal.pone.0175054
24. Sheng Z, Jia X, Kang M. Statin use and risk of Parkinson's disease: A meta-analysis. *Behav Brain Res*. 2016;309:29-34. doi:10.1016/j.bbr.2016.04.046
25. Undela K, Gudala K, Malla S, Bansal D. Statin use and risk of Parkinson's disease: a meta-analysis of observational studies. *J Neurol*. 2013;260(1):158-165. doi:10.1007/s00415-012-6606-3
26. Nicoletti A, Vasta R, Mostile G, et al. Head trauma and Parkinson's disease: results from an Italian case-control study. *Neurol Sci*. 2017;38(10):1835-1839. doi:10.1007/s10072-017-3076-5
27. Harris MA, Shen H, Marion SA, Tsui JKC, Teschke K. Head injuries and Parkinson's disease in a case-control study. *Occup Environ Med*. 2013;70(12):839-844. doi:10.1136/oemed-2013-101444
28. Jafari S, Etminan M, Aminzadeh F, Samii A. Head injury and risk of Parkinson disease: a systematic review and meta-analysis. *Mov Disord*. 2013;28(9):1222-1229. doi:10.1002/mds.25458
29. Kenborg L, Rugbjerg K, Lee PC, et al. Head injury and risk for Parkinson disease: results from a Danish case-control study. *Neurology*. 2015;84(11):1098-1103. doi:10.1212/WNL.0000000000001362
30. Fang F, Chen H, Feldman AL, Kamel F, Ye W, Wirdefeldt K. Head injury and Parkinson's disease: a population-based study. *Mov Disord*. 2012;27(13):1632-1635. doi:10.1002/mds.25143
31. Spangenberg S, Hannerz H, Tüchsen F, Mikkelsen KL. A nationwide population study of severe head injury and Parkinson's disease. *Parkinsonism Relat Disord*. 2009;15(1):12-14. doi:10.1016/j.parkreldis.2008.02.004
32. Lee PC, Bordelon Y, Bronstein J, Sinsheimer JS, Farrer M, Ritz B. Head injury, α -synuclein genetic variability and Parkinson's disease. *Eur J Neurol*. 2015;22(5):874-878. doi:10.1111/ene.12585
33. Gao J, Liu R, Zhao E, et al. Head injury, potential interaction with genes, and risk for Parkinson's disease. *Parkinsonism Relat Disord*. 2015;21(3):292-296. doi:10.1016/j.parkreldis.2014.12.033
34. Goldman SM, Kamel F, Ross GW, et al. Head injury, α -synuclein Rep1, and Parkinson's disease. *Ann Neurol*. 2012;71(1):40-48. doi:10.1002/ana.22499

35. Rugbjerg K, Ritz B, Korbo L, Martinussen N, Olsen JH. Risk of Parkinson's disease after hospital contact for head injury: population based case-control study. *BMJ*. 2008;337:a2494. doi:10.1136/bmj.a2494
36. Taylor KM, Saint-Hilaire MH, Sudarsky L, et al. Head injury at early ages is associated with risk of Parkinson's disease. *Parkinsonism Relat Disord*. 2016;23:57-61. doi:10.1016/j.parkreldis.2015.12.005
37. Cheng CM, Wu YH, Tsai SJ, et al. Risk of developing Parkinson's disease among patients with asthma: a nationwide longitudinal study. *Allergy*. 2015;70(12):1605-1612. doi:10.1111/all.12758
38. Li CH, Chen WC, Liao WC, et al. The association between chronic obstructive pulmonary disease and Parkinson's disease: a nationwide population-based retrospective cohort study. *QJM*. 2015;108(1):39-45. doi:10.1093/qjmed/hcu136
39. Mittal S, Bjørnevik K, Im DS, et al. β 2-Adrenoreceptor is a regulator of the α -synuclein gene driving risk of Parkinson's disease. *Science*. 2017;357(6354):891-898. doi:10.1126/science.aaf3934
40. Gronich N, Abernethy DR, Auriel E, Lavi I, Rennert G, Saliba W. β 2-adrenoceptor agonists and antagonists and risk of Parkinson's disease. *Mov Disord*. 2018;33(9):1465-1471. doi:10.1002/mds.108
41. Chen CL, Wang SY, Chen TC, Chuang CS. Association between β 2-Adrenoreceptor Medications and Risk of Parkinson's Disease: A Meta-Analysis. *Medicina (Kaunas)*. 2021;57(10):1006. doi:10.3390/medicina57101006
42. Gustafsson H, Nordström A, Nordström P. Depression and subsequent risk of Parkinson disease: A nationwide cohort study. *Neurology*. 2015;84(24):2422-2429. doi:10.1212/WNL.0000000000001684
43. Zis P, Erro R, Walton CC, Sauerbier A, Chaudhuri KR. The range and nature of non-motor symptoms in drug-naive Parkinson's disease patients: a state-of-the-art systematic review. *NPJ Parkinsons Dis*. 2015;1:15013. doi:10.1038/npjparkd.2015.13
44. Schrag A, Taddei RN. Depression and Anxiety in Parkinson's Disease. *Int Rev Neurobiol*. 2017;133:623-655. doi:10.1016/bs.irn.2017.05.024
45. Gagne JJ, Power MC. Anti-inflammatory drugs and risk of Parkinson disease: a meta-analysis. *Neurology*. 2010;74(12):995-1002. doi:10.1212/WNL.0b013e3181d5a4a3
46. Ren L, Yi J, Yang J, Li P, Cheng X, Mao P. Nonsteroidal anti-inflammatory drugs use and risk of Parkinson disease: A dose-response meta-analysis. *Medicine (Baltimore)*. 2018;97(37):e12172. doi:10.1097/MD.00000000000012172

47. Asanuma M, Miyazaki I. Nonsteroidal anti-inflammatory drugs in experimental parkinsonian models and Parkinson's disease. *Curr Pharm Des.* 2008;14(14):1428-1434. doi:10.2174/138161208784480153
48. Goesling J, Clauw DJ, Hassett AL. Pain and depression: an integrative review of neurobiological and psychological factors. *Curr Psychiatry Rep.* 2013;15(12):421. doi:10.1007/s11920-013-0421-0
49. Payami H. The emerging science of precision medicine and pharmacogenomics for Parkinson's disease. *Mov Disord.* 2017;32(8):1139-1146. doi:10.1002/mds.27099

Figure Legends

Figure 1: Schematic demonstrating the variables examined in this study. Variables span domains of past medical history, medication and supplement use, social history, and environmental exposures. They capture functions of many organ systems.

Figure 2: Heat map demonstrating correlations between the individual environmental variables. A. Pearson correlation matrix. B. Several sets of variables are positively correlated, forming clusters of biologically related phenomena. The largest clusters in our dataset include a linked sleep and psychiatric factor cluster, a metabolic syndrome cluster, and a vitamin use cluster. P-values indicating the significance of the correlation are shown.

Figure 3: Results of logistic regression with backward elimination. Male sex, coronary artery disease, depression, anxiety, restless leg syndrome, head trauma, ibuprofen, co-enzyme Q10, and vitamin D were over-represented in PD. Asthma/chronic obstructive pulmonary disease (COPD), naproxen, ezetimibe use, and smoking as under-represented in PD. Odds ratio with 95% confidence limit for each significant variable in the logistic regression is shown. For the continuous variable of age, the odds ratio is listed for incremental increases of 1 year, 5 years, or 10 years.

Figure 4: Among the variables identified in the logistic regression, small but statistically significant correlations exist. P-values <0.01 are shown within the corresponding squares.

Figure 5: Odds ratio for pairwise combinations of variables enriched in PD. A. Distribution of variables positively associated with PD across the HC and PD groups. In the HC group, the majority of subjects are positive for none of the variables associated with PD. In the PD group, the majority of subjects are positive for 1 or more variables associated with PD. B. Distribution of pairwise combinations of variables positively associated with PD among the PD group. C. Odds ratio and confidence limits for pairwise combinations of variables positively associated with PD. Note: For some combinations, OR cannot be calculated because no HC subjects were positive for that pairwise combination. D. Prevalence of depression increases with increasing number of positive variables in PD subjects.

Figure 1

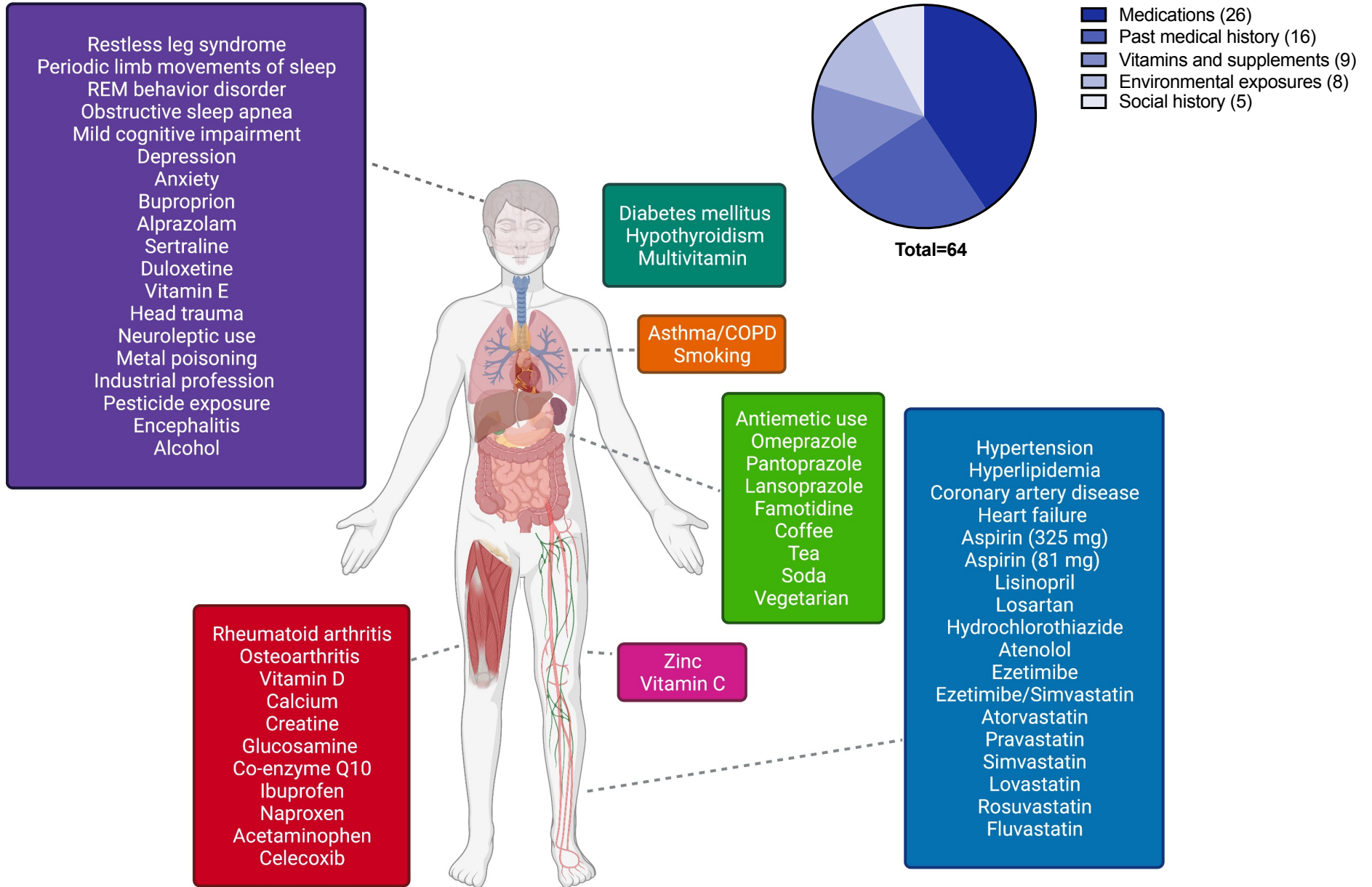


Figure 2A

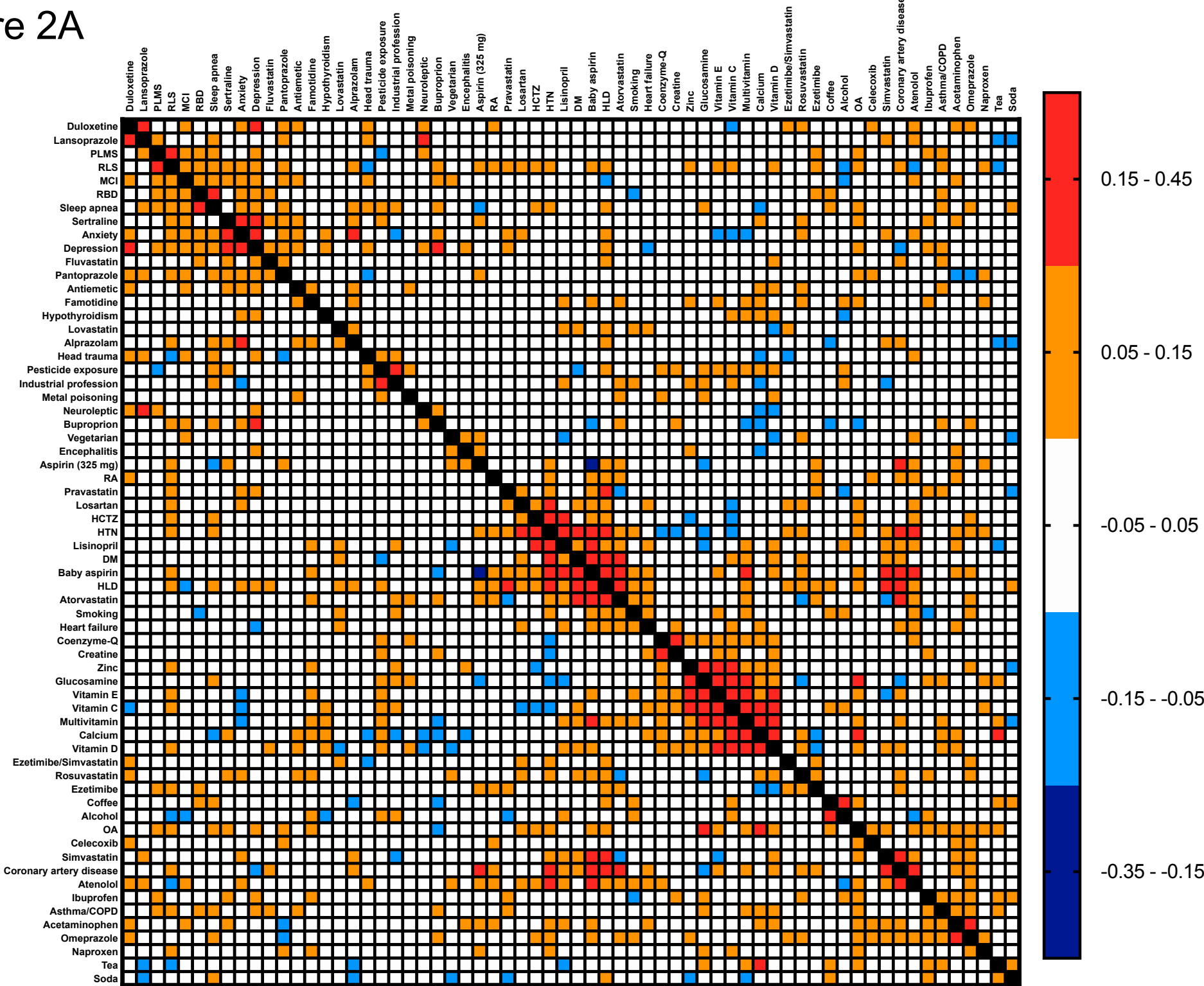


Figure 2B

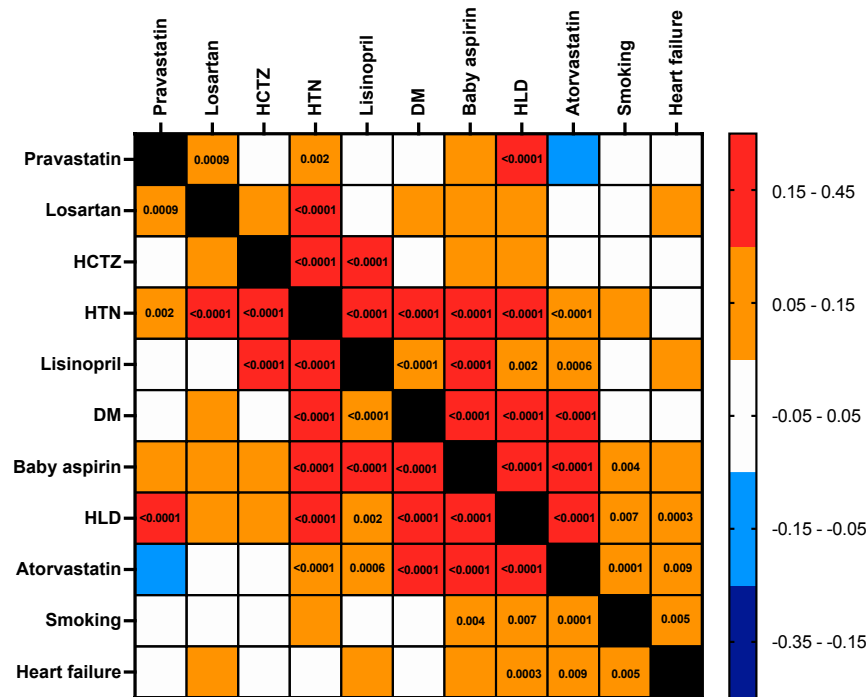
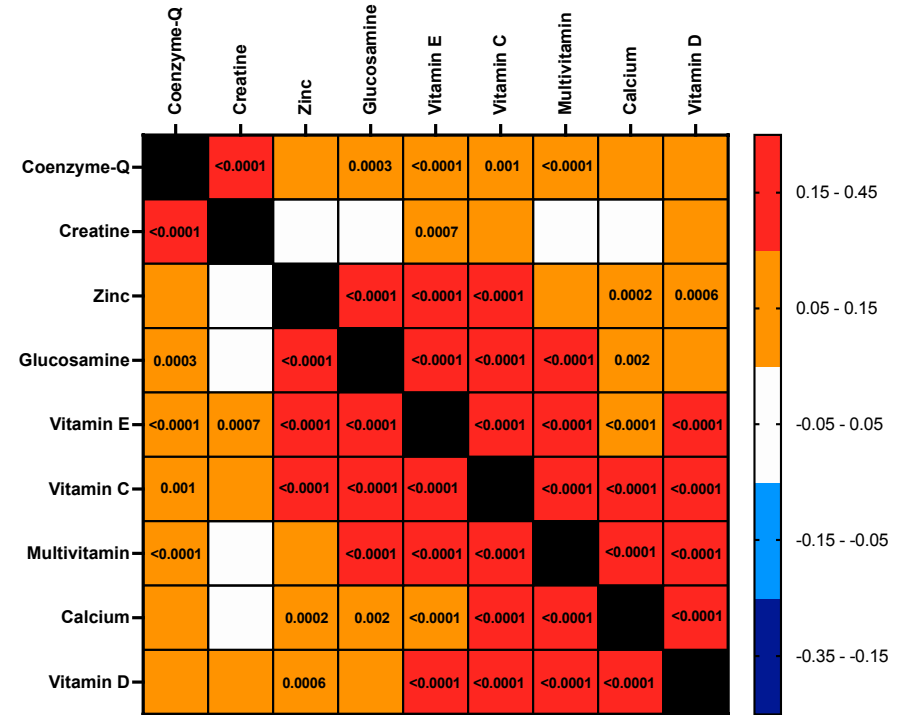
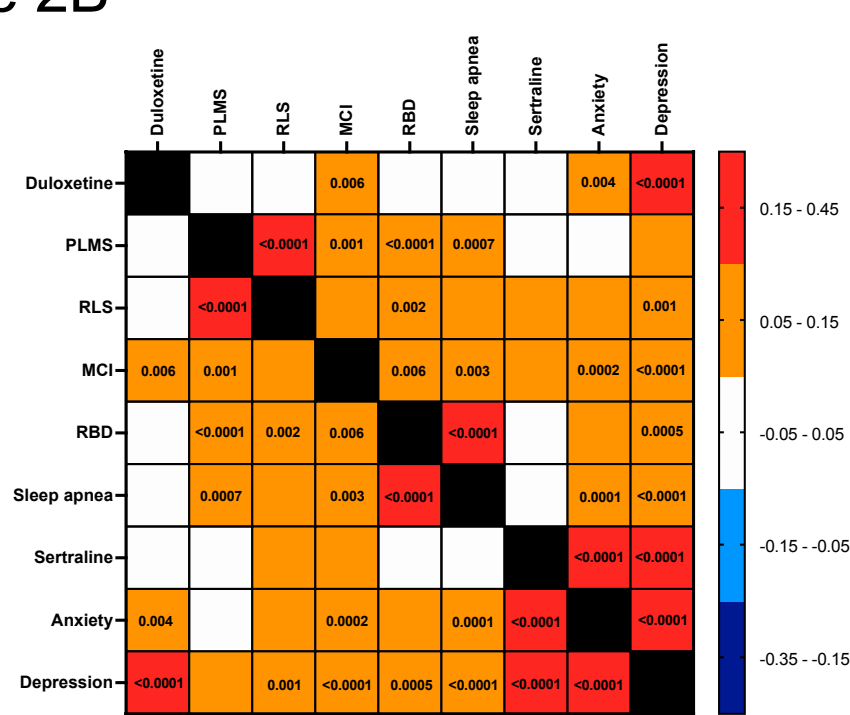
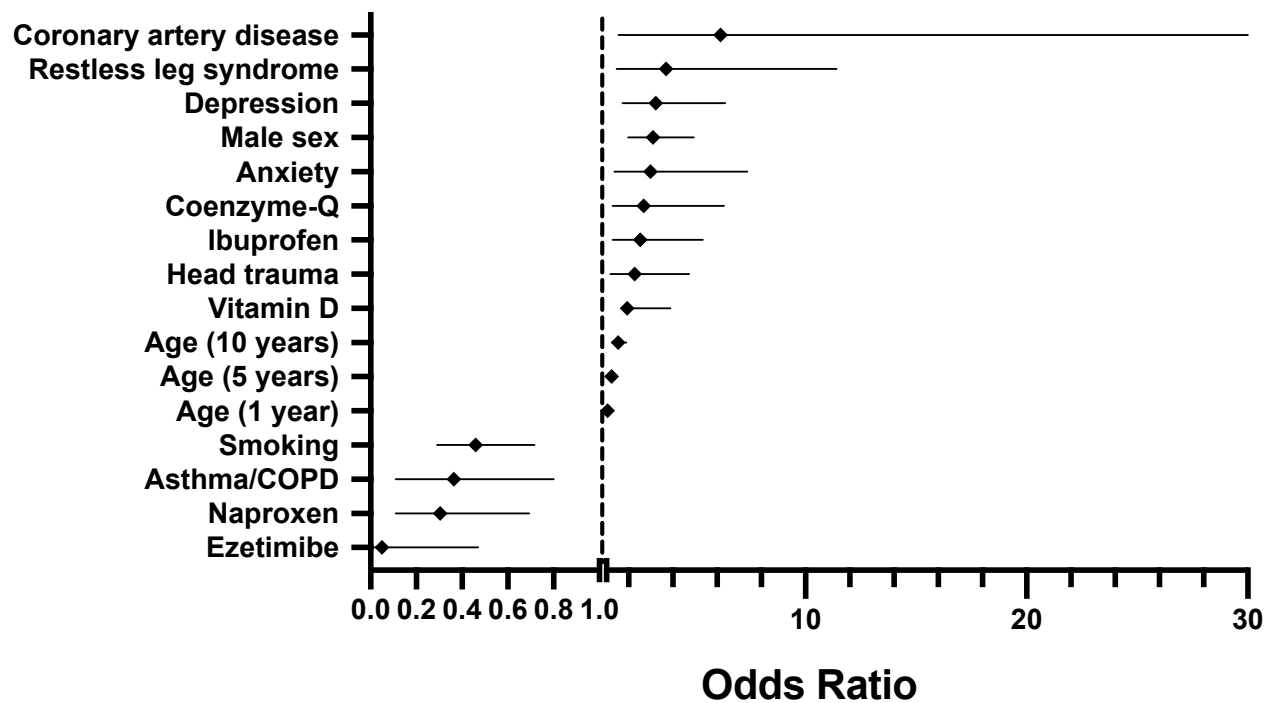
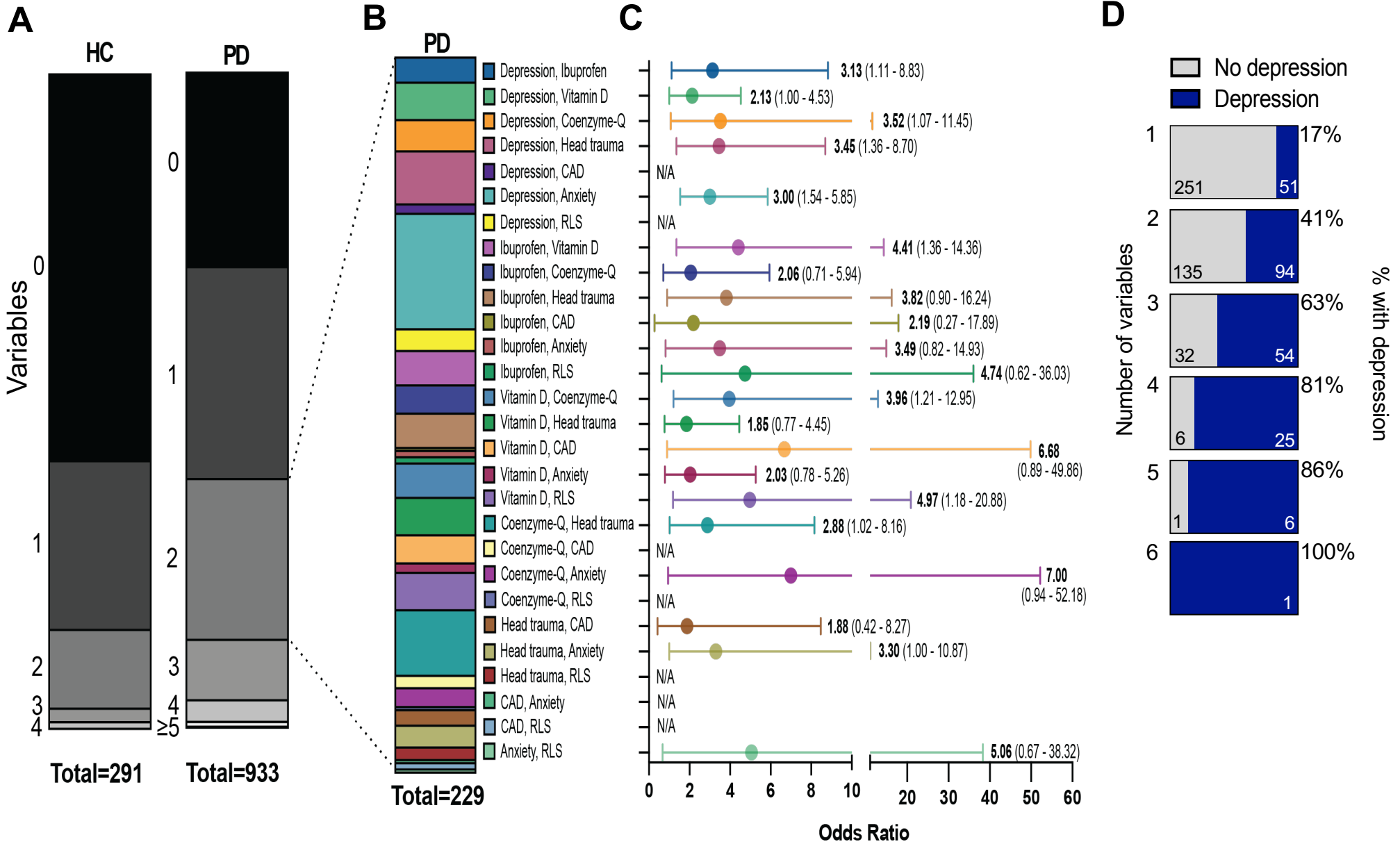


Figure 3



Variable	OR	Lower limit	Upper limit
CAD	6.15	1.53	49.12
RLS	3.69	1.46	11.41
Depression	3.23	1.73	6.37
Male	3.11	1.97	4.94
Anxiety	2.99	1.36	7.36
Co-Q	2.68	1.27	6.31
Ibuprofen	2.52	1.27	5.36
Head trauma	2.27	1.18	4.73
Vit D	1.94	1.86	3.89
Age (10)	1.53	1.25	1.89
Age (5)	1.24	1.12	1.38
Age (1)	1.04	1.02	1.07
Smoking	0.46	0.29	0.72
Asthma/COPD	0.36	0.11	0.80
Naproxen	0.30	0.11	0.69
Ezetimibe	0.05	0.004	0.47

Figure 5



Tables

Table 1: Demographics of HBS. Statistical differences in demographic data (age, sex, and race) between the cases and controls were determined using a Satterwaithe t-test, Chi-square test and Fisher’s exact test, as appropriate. UPDRS = Unified Parkinson’s Disease Rating Scale. MMSE = Mini-Mental State Exam.

Table 1: Demographics in the Harvard Biomarkers Study			
Characteristic	Healthy Controls (HC)	Parkinson’s disease (PD)	p-value
Sex			<0.0001
Male	125 (43%)	620 (66.5%)	
Female	166 (57%)	312 (33.5%)	
Age at enrollment (years)	63.5 ±11.7	66.1 ± 9.9	0.0006
Race			
Asian	2 (0.7%)	9 (0.7%)	1.000
Black/African American	13 (4.9%)	9 (0.7%)	<0.0001
Hawaiian/Pacific	1 (0.3%)	0	0.1920
More than one race	3 (1%)	3 (0.2%)	0.0888
Unknown	4 (1.4%)	38 (3.1%)	0.1154
White	268 (92.1%)	1166 (95.2%)	0.0432
Ethnicity			
Hispanic	8 (2.7%)	17 (1.4%)	0.1211
Non-Hispanic	277 (95.2%)	1168 (95.3%)	0.8779
Unknown	6 (2.1%)	40 (3.3%)	0.3447
Average education level (years, S.D.)	15.0 ± 1.77	15.2 ± 1.76	0.159
Age at onset	N/A	62.3 ± 10.4	N/A
Disease duration	N/A	3.8 ± 4.4	N/A
Hoehn and Yahr stage	N/A	2.1 ± 0.64	N/A
UPDRS motor subscale	N/A	19.8 ± 10.2	N/A
UPDRS total	N/A	33.2 ± 15.1	N/A
MMSE	28.99 ± 1.26	28.21 ± 2.30	<0.0001

Table 2: Results of multivariate logistic regression after backward elimination. P-value, odds ratio, 95% confidence limit, and prevalence in HC versus PD for the 13 retained variables are shown.

No.	Feature	P-value (adjusted for all other variables)	Odds ratio estimate	95% C.I.	Healthy Controls	Parkinson's disease
1	Age (1 year increase)	<0.0001	1.044	1.022 - 1.066		
2	Age (5 year increase)		1.237	1.117 - 1.375		
3	Age (10 year increase)		1.531	1.248 - 1.890		
2	Male gender	<0.0001	3.107	1.972 - 4.944		
3	Coronary artery disease	0.0313	6.154	1.539 - 49.123	9 (3.21%)	66 (7.42%)
4	Depression	0.0004	3.225	1.726 - 6.359	34 (12.2%)	229 (25.7%)
5	Anxiety	0.0101	2.992	1.360 - 7.364	18 (6.5%)	140 (15.8%)
6	Asthma/COPD	0.0109	0.364	0.168 - 0.801	23 (8.2%)	53 (6.0%)
7	Restless leg syndrome	0.0112	3.691	1.464 - 11.410	5 (3.0%)	75 (12.6%)
8	Ibuprofen	0.0113	2.520	1.271 - 5.357	28 (9.9%)	127 (14.2%)
9	Naproxen	0.0250	0.304	0.109 - 0.888	12 (4.2%)	26 (2.9%)
10	Ezetimibe	0.0107	0.050	0.004 - 0.470	3 (1.9%)	3 (0.52%)
11	Head trauma	0.0202	2.269	1.175 - 4.732	25 (9.1%)	176 (19.8%)
12	Coenzyme Q	0.0149	2.681	1.269 - 6.308	22 (7.7%)	170 (19.0%)
13	Vitamin D	0.0092	1.937	1.856 - 3.888	34 (23.6%)	202 (40.3%)
14	Smoking	0.0007	0.459	0.291-0.717	142 (50.1%)	365 (41.1%)