# Mitochondrial and autophagy-lysosomal pathway polygenic risk scores predict Parkinson's disease

Mohammad Dehestani [a,b,*], Hui Liu [a,b], Ashwin Ashok Kumar Sreelatha [c], Claudia Schulte [a,b], Vikas Bansal [b], Thomas Gasser [a,b]

[a] *Department of Neurodegenerative Disease, Her tie Institute for Clinical Brain Research, University of Tübingen, Germany*
[b] *German Center for Neurodegenerative Diseases DZNE, Tübingen, Germany*
[c] *Centre for Genetic Epidemiology, Institute for Clinical Epidemiology and Functional Biometry, University of Tübingen, Germany*

## ARTICLE INFO

## ABSTRACT

Polygenic Risk Scores (PRS), which allow assessing an individuals' genetic risk for a complex disease, are calculated as the weighted number of genetic risk alleles in an individual's genome, with the risk alleles and their weights typically derived from the results of genome-wide association studies (GWAS). Among a wide range of applications, PRS can be used to identify at-risk individuals and select them for further clinical follow-up. Pathway PRS are genetic scores based on single nucleotide polymorphisms (SNPs) assigned to genes involved in major disease pathways. The aim of this study is to assess the predictive utility of PRS models constructed based on SNPs corresponding to two cardinal pathways in Parkinson's disease (PD) including mitochondrial PRS (Mito PRS) and autophagy-lysosomal PRS (ALP PRS). PRS models were constructed using the clumping-and-thresholding method in a German population as prediction dataset that included 371 cases and 249 controls, using SNPs discovered by the most recent PD-GWAS.

We showed that these pathway PRS significantly predict the PD status. Furthermore, we demonstrated that Mito PRS are significantly associated with later age of onset in PD patients.

Our results may add to the accumulating evidence for the contribution of mitochondrial and autophagy-lysosomal pathways to PD risk and facilitate biologically relevant risk stratification of PD patients.

## 1. Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disease. Sporadic PD is considered a complex disorder in which the genetic component, represented by numerous risk and protective alleles, interacts with environmental and ageing factors to contribute to the formation and propagation of alpha-synuclein aggregates and ultimately to neurodegeneration (Gasser, 2007). This interplay of genetic, environmental and ageing factors, the components of which vary from person to person, could lead to differential susceptibility to disease.

Over the past decade, genome-wide association studies (GWAS) in PD have uncovered a large number of genetic risk variants, single nucleotide polymorphisms (SNPs), that influence the likelihood of developing PD, age at onset, and disease progression (Fung et al., 2006; Chang et al., 2017; Nalls et al., 2019). Polygenic risk scores (PRS) are new tools that summarize the genetic risk by quantifying the combined effect of risk SNPs in each individual. In other words, PRS as a weighted

sum of all disease-relevant genetic variants detected by sufficiently powerful large GWAS could serve as a predictor of disease risk and also provide mechanistic insights into the disease or trait of interest (Dudbridge, 2013).The most common application of PRS is to stratify populations and identify individuals at high genetic risk (Duncan et al., 2019). When the biological basis of a disease is relatively discovered in terms of the major pathways involved, it is possible to construct PRS using SNPs that enriched in a particular disease-related pathway.

In PD, two seminal disease pathways, the mitochondrial and autophagy-lysosomal pathways (ALP) have been highlighted to be of particular importance. Accumulating evidence suggests that mitochondrial dysfunction is an essential component of the PD etiology (Park et al., 2018). Also, genetic studies focusing on monogenic forms of PD have provided insight for the involvement of mitochondrial dysfunction (Antony et al., 2013). Likewise, the ALP plays an important role as a cellular housekeeping mechanism in maintaining the effective turnover of proteins and its dysfunction has been implicated in PD (Pan et al.,

2008). Any impairment of these tightly regulated pathways could lead to degeneration of dopaminergic neurons (Finkbeiner, 2020; Hou et al., 2020). The dysfunction of these pathways may be due partly to the presence of risk variants in nuclear genes encoding their components. In other words, Parkinson's disease cases are more likely to be carriers of risk variants in genes involved in mitochondrial and autophagic-lysosomal pathways. With this in mind, pathway PRS allow for the identification and follow-up of patients at highest risk with respect to specific pathways.

The aim of this study is to assess the predictive value of two pathway PRS including Mito and ALP PRS in predicting PD as well as relevant PD clinical outcomes including age at onset (AAO), cognitive performance (represented by Montreal Cognitive Assessment or MoCA) and motor progression (represented by Unified Parkinson's Disease Rating Scale III or UPDRS III).

## 2. Material and methods

### 2.1. Base data

The GWAS summary statistics from the largest and most recent PD metaGWAS was used as the base dataset. This meta-analysis included 7.8 million SNPs in 37.7 K cases, 18.6 K UK Biobank proxy cases, and 1.4 million controls. The 23&Me data was excluded from this summary statistics (Nalls et al., 2019).

### 2.2. Target data

For the prediction and validation, imputed genotypes of 686 cases and 544 controls from Tuebingen, which were not included in the base data, were used. Samples were primarily genotyped on Neurochip (Blauwendraat et al., 2017) as part of the Comprehensive Unbiased Risk Factor Assessment for Genetics and Environment in Parkinson's Disease (COURAGE-PD) project. Imputation was done in our cloned Michigan Imputation Server (MIS) (https://193.196.20.138:8080/) at the DENBI cloud (https://denbi.uni-tuebingen.). The Haplotype Reference Consortium Release 1.1 (HRC) data usage request was approved by the Sanger Institute (Dataset ID: EGAD00001002729). The datasets were prepared under the reference panel criteria for the MIS (https://imputationserver.). We have used the HRC/1000G imputation preparation and checking tool (https://www.well.ox.ac.uk/~) to check for Ref/Alt allele assignments, incorrect strands, deviation from allele frequency, and palindromic SNPs. Later, the chip genotype dataset after quality control (QC) was phased using Eagle v2.4 in our MIS. Imputation of autosomal variants was performed separately for each dataset through the MIS using the HRC reference panel and the GRCh37/hg19 assembly. Finally, imputed data were subjected to hard-calling with a filter of 0.9 on PLINK2 (Purcell and Chang, n.d.) and then converted to binary genotypes (bed, bim, fam format) to be used for further analyses. The dataset was also randomly split into a 70 % part for prediction and a 30 % part for validation. Descriptive cohort statistics are provided in Supplementary Data 3.

### 2.3. SNP selection for pathway PRS models: Mitochondrial (Mito) and Autophagy-Lysosomal (ALP) models

Using all SNPs from the base summary statistics filtered with a Minor Allele Frequency (MAF) of 0.01, we performed a comprehensive annotation using the Region Annotation function in Annovar to find all genes corresponding to these SNPs in the whole genome and then generated the GWAS gene list. Next, to determine the proportion of GWAS genes implicated in mitochondrial pathways, we examined the overlap between these annotated genes and the full list of human genes with mitochondrial evidence from two databases that happened to overlap approximately 70 %. The first was Human MitoCarta2.0 (N = 1158 genes), the most comprehensive list of human repository genes

containing nuclear and mtDNA genes encoding proteins with clear evidence of mitochondrial localization (Calvo et al., 2016). We also used a gene list from the Billingsley et al. study (N = 1639 genes), which includes mitochondria-related genes as well as genes involved in mitochondrial diseases (Billingsley et al., 2019).

We compiled all SNPs assigned to these genes from the annotation results and then retrieved their information, including MAF, beta, and p-value from the base data summary statistics. By filtering for p-value of 0.05 and MAF of 0.01, we obtained 452,560 SNPs assigned to mitochondrial genes which further reduced to 1356 SNPs after clumping procedure for the construction of the Mito PRS model.

In a second set of analysis, we used all GWAS genes designated as "lysosomal" in the Gene Ontology (GO:0006914), in addition to all genes corresponding to the autophagy pathway according to the HADb database (http://autophagy.lu/index.html). For 465 autophagy genes and 232 lysosomal genes combined, we obtained 198,464 SNPs from the Base summary statistics. After filtering for a p-value of 0.05, a MAF of 0.01 and clumping, 546 SNPs remained to be used for the ALP PRS model.

### 2.4. PRS control models

In a separate analysis, we performed a sanity check to further control the contingency of selected SNPs in our PRS models and their contribution to PD risk. We used the SNPs corresponding to 200 genes involved in myogenesis as a PD-independent pathway from MSigDB (https://www.gsea-msigdb.org/gsea/msigdb/) (Liberzon et al., 2015) as well as a positive control from SNPs corresponding to 1028 genes involved in the inflammation as a PD related pathway from Loza et al. study (Loza et al., 2007). Also, to further control for the actual contribution of PD SNPs to the prediction and risk stratification, we used SNPs of another brain-related disorder, namely attention deficit hyperactivity disorder (ADHD), from the GWAS summary statistics of Demontis et al. (2019) to construct another PRS model and applied it to our prediction population. In addition, we used only hit GWAS SNPs (N = 90) to construct general PD PRS so that we could compare them with the pathway PRS in our dataset.

### 2.5. PRS calculation method

We used the R package PRSice2 to construct the PRS models (Choi and O'Reilly, 2019). The method implemented in PRSice2 is clumping plus thresholding (C + T). The scores are defined as the sum of the allele counts weighted by the estimated effect sizes (GWAS beta). In C + T, variants are first clumped (C) within a specified LD window ($\pm 1$ Mb), leaving only variants that were weakly correlated with each other. Clumping iteratively selects the most significant variant, calculates the correlation between this index variant and nearby variants within a certain LD, and removes all nearby variants correlated with this index variant beyond a certain window i.e. r2 of 0.5. Thresholding (T) consists of removing variants with a p-value greater than a chosen significance level. Nevertheless, by choosing a thresholding p-value of 0.05 in the interest of avoiding the inclusion of non-significant SNPs, a certain range of GWAS p-value thresholds (from 5e-08 to 0.05) was chosen. In other words, all SNPs with a p-value below 0.05 from the LD-clumped list were included. Since C + T is basically a heuristic method, various tuning parameters, i.e. clump r2, clump window, and p-value thresholds were tested to find the best setting required to maximize the predictive power for each PRS model (Choi et al., 2020).

### 2.6. PRS model fit and predictive accuracy

The predictive value of a PRS model is measured by its variance explained and its predictive accuracy. The explained variance is the PRS model fit or incremental $R^2$ obtained by subtracting the $R^2$ of the full model from the $R^2$ of the null model. The null model is a logistic

regression model that measures the power of covariates including age, sex, and principal genetic components (PC1–4) in the prediction of PD status, whereas the full model adds the PRS to the null model, thereby isolating the additive influence of the PRS on risk prediction. The $R^2$ was also adjusted for an estimated PD prevalence of 0.005 on the liability scale (Bandres-Ciga et al., 2020). The PRS model corresponding to the GWAS p-value threshold with the highest $R^2$ was selected as the best model. The predictive accuracy of the PRS models was assessed using the Area Under the Receiver Operating Curve (AUROC) analysis. The AUC is interpreted as the probability that a case ranks higher than the control, and by analogy, the higher the AUC, the better a PRS model can discriminate between cases and controls (Huang and Ling, 2005).

### 2.7. PRS model validation

In addition to performing the permutation test that serves as a complementary method to account for overfitting and yields an empirical p-value for each PRS model (Choi et al., 2020), we also performed cross-validation or out-of-sample validation by measuring the variance explained i.e. model fit $R^2$ and predictive accuracy i.e. AUC of PRS models in the validation dataset comprising 210 cases and 159 controls.

### 2.8. Correlation of PRS models with PD clinical outcomes

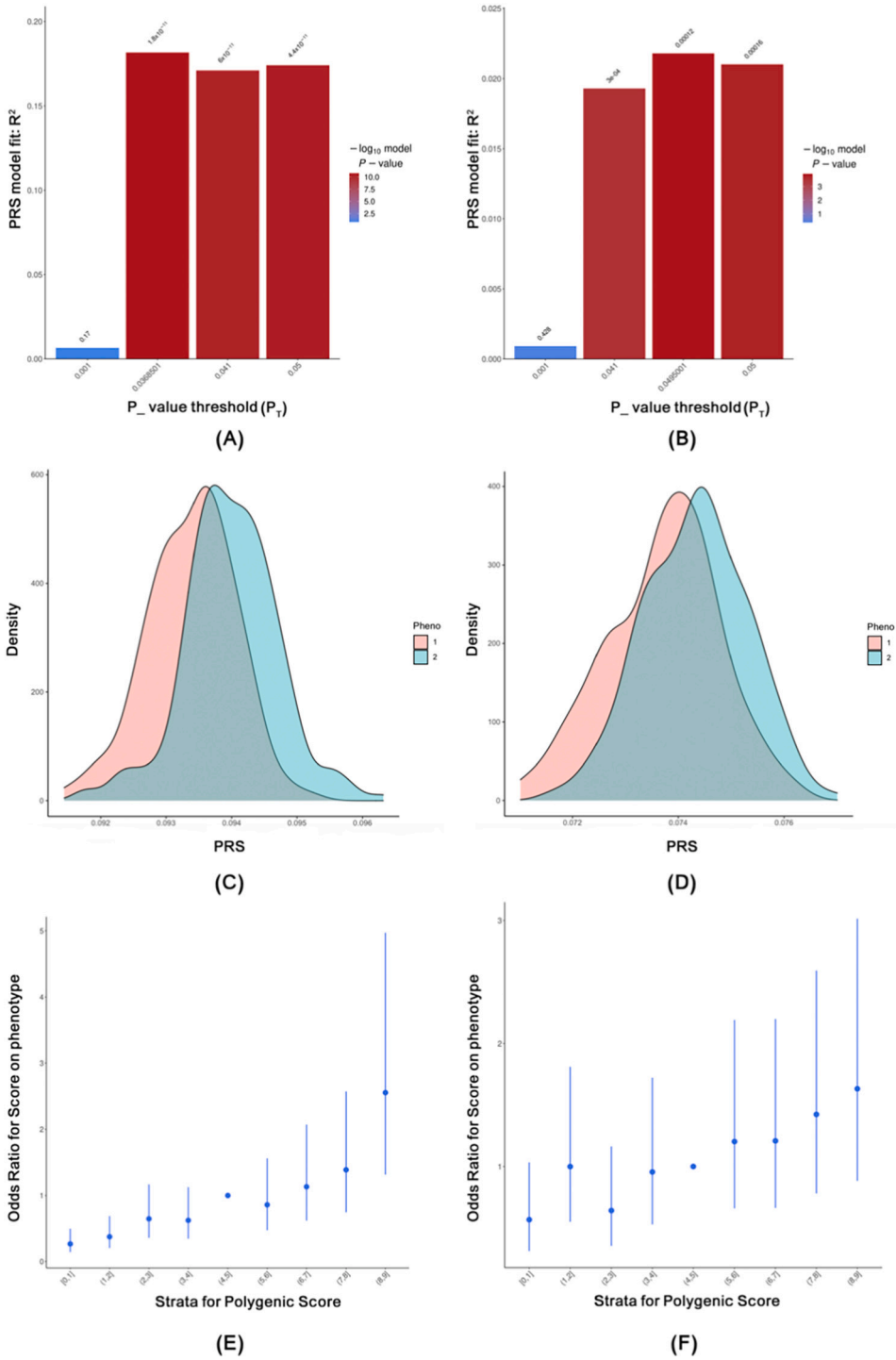To explore the prediction of the pathway PRS models for PD clinical



**Fig. 1.** The prediction barplots generated by PRSice2 display different GWAS p-value thresholds ranging up to 0.05 on the X-axis and model-fit $R^2$ on the Y-axis along with its p-value on the top of each bar for Mito PRS (A) and ALP PRS (B). Prediction density plots generated by ggplot2 (Wickham, 2016) depict the discrimination ability between cases (Pheno = 2) and controls (Pheno = 1) by Mito PRS (C) and ALP PRS (d). Also, the strata plots generated by PRSice2 illustrate the fold change in genetic risk across the entire population by plotting the odds ratio of risk on the Y-axis and the polygenic score quantiles on the X-axis for Mito-PRS (E) and ALP-PRS (F).

course, outcomes including AAO, cognitive performance, and motor progression (measured with the MoCA and UPDRS-III scales, respectively) were selected. Statistical analysis was performed using the Spearman correlation test and linear regression model with covariates including age, sex and disease duration for the cross-sectional outcomes, ie, AAO, and the mixed-effects linear regression model for the longitudinal outcomes, including MoCA and UPDRS III that have repeated measures. In the mixed-effects linear regression models, covariates including age, sex and disease duration were used, and the random-effects term was grouped by individuals to account for repeated measures.

## 3. Results

### 3.1. PRS model fit and predictive accuracy

Both PRS models significantly predicted PD status (Fig. 1). The Mito model showed the best prediction at GWAS p-value threshold of 0.03 ($R^2$: 0.10, p-value = 1.8e-11, empirical p = 9.999e-05), and the ALP model at GWAS p-value threshold of 0.04 ($R^2$: 0.02, p-value = 0.00012, empirical p = 0.0013). These PRS were then carried forward to further analyses. The measure of predictive accuracy, i.e., AUC showed appropriate performance, including 67 % for the Mito PRS and 58 % for the ALP PRS (detailed results in Table 1).

### 3.2. PRS control models

The negative control PRS model, which was based on myogenesis-related SNPs, behaved as expected, showing very low model fit. ($R^2$ = 0.0045, p-value = 0.03). The positive control model, a PRS model based on a fraction of the inflammation-related SNPs (equal to the number of SNPs in the negative control model to avoid SNP count bias), also showed sufficient PRS model fit ($R^2$ = 0.027, p-value = 2.3e-7). The PRS results for ADHD also showed no significant prediction ($R^2$ = 0.001, p-value = 0.2) when applied to our prediction population. The attempt of constructing PRS model based on only hit GWAS SNPs resulted in modest prediction (R2 = 0.04, p-value = 9.7e-8). Detailed results on control PRS models are provided in the supplementary data 1.

### 3.3. PRS models validation

Variance explained, empirical p-values, and predictive accuracy showed significant validation of both PRS models in the validation dataset (Table 1). Detailed validation results for each PRS model can be found in the supplementary Data 2.

### 3.4. Correlation of PRS models with PD clinical outcomes

Among all, only significant correlation was observed for Mito PRS and AAO tested by both methods, including Spearman test without covariates and linear regression with covariates, including age, sex and disease duration. In addition, there was no statistically significant

prediction in both PRS models for longitudinal outcomes including MoCA and UPDRS III. (Table 2).

## 4. Discussion

Since the advent of GWA studies that have been increasingly identifying risk SNPs associated with complex diseases, polygenic scoring based on these risk variants has gained traction as a method for predicting disease status and for identifying high-risk individuals. It also appears that increasing the number of SNPs obtained from GWAS facilitates the curation of certain SNPs associated with genes involved in specific disease pathways and in this way allows the construction of pathway PRS. These PRS can potentially indicate and quantify the contribution of a pathway to a disease in each individual (Lewis and Vassos, 2020).

In the design of any PRS study, regardless of its subsequent utility, the first fundamental question to address is how to select SNPs from base summary statistics data and what weights to apply. Conventionally, SNPs reaching the universal GWAS *p*-value threshold (5e-08), have been used in the PRS design, and GWAS betas (or odds ratios) were used as weights. However, to achieve more power, we can relax this threshold and move from 5e-8 to a lower threshold to allow the inclusion of sub-significant SNPs, or even move to a genome-wide selection of millions of SNPs to increase the predictive value. However, this gain in predictive power could be at the cost of increasing the probability of

**Table 2**
Statistical analysis of the different clinical outcomes for each PRS model.

| PRS model | clinical outcome | N | Statistical test | Coefficient * | P-value |
|---|---|---|---|---|---|
| Mito | AAO | 305 | Spearman correlation (without covariates) | **r = 0.12** | **0.03** |
| | | | Linear regression | **$R^2$ = 0.05** | **0.02** |
| | UPDRS III | 305 | Mixed effects regression | $R^2$ = 0.022/ 0.509 | 0.35 |
| | MoCA | 305 | Mixed effects regression | $R^2$ = 0.083/ 0.706 | 0.08 |
| ALP | AAO | 305 | Spearman correlation (without covariates) | r = 0.03 | 0.49 |
| | | | Linear regression | $R^2$ = 0.04 | 0.36 |
| | UPDRS III | 305 | Mixed effects regression | $R^2$ = 0.021/ 0.510 | 0.67 |
| | MoCA | 305 | Mixed effects regression | $R^2$ = 0.078/ 0.706 | 0.52 |

*: Coefficient means the Spearman coefficient (r) for the Spearman test, the adjusted $R^2$ for the linear regression model, and the marginal $R^2$/conditional $R^2$ for the mixed-effect regression model. Marginal $R^2$ provides the variance explained only by fixed effects and conditional $R^2$ provides the variance explained by the entire model, i.e., both fixed effects and random effects. Bold values indicate statistical significant results (p-value < 0.05).

**Table 1**
Summary results of the PRS models performance for both prediction and validation including model fit $R^2$ on the liability scale with PD prevalence of 0.005, Empirical P-values, and Predictive accuracy (AUC).

| PRS Model | Prediction 371 cases and 249 controls | | | | Validation 210 cases and 159 controls | | | |
|---|---|---|---|---|---|---|---|---|
| | N of clumped SNPs | Empirical p-value | Adjusted incremental $R^2$ | AUC (95 % CI) | N of clumped SNPs | Empirical p-value | Adjusted incremental $R^2$ | AUC (95 % CI) |
| Mito PRS | 1356 | 9.999e-05 | 0.10 | 0.67 (0.595, 0.6609) | 1049 | 9.999e-05 | 0.18 | 0.73 (0.6269, 0.7255) |
| ALP PRS | 538 | 0.0013 | 0.02 | 0.58 (0.5334, 0.601) | 533 | 9.999e-05 | 0.09 | 0.65 (0.5509, 0.654) |

overestimation in the final PRS model.

The main objective of this study was to explore the potential use of pathway PRS as a predictor of PD status and clinical outcomes by looking into two major PD pathways, including mitochondrial and autophagy-lysosomal. Therefore, two PRS models were created using different SNP sets corresponding to these two pathways. The predictive value was measured by variance explained or model fit $R^2$, i.e., the ability of the PRS model to explain actual variation in individuals by identifying the best scores over a range of GWAS $p$-value thresholds. Also, predictive accuracy was measured by AUROC analysis which is interpreted as the probability that a case ranks higher than the control, and by analogy, the higher the AUC, the better a PRS model can discriminate between cases and controls. Since a standard method for validating PRS models is to apply them to an independent cohort and then to measure their predictive value, the AUC analysis was also performed for the validation population. Also a PRS model based on GWAS hit SNPs ($N = 90$) was constructed to be able to compare it with pathway PRS models. Both pathway PRS models showed significant prediction for PD status ($R^2$ of 0.10 for Mito model and $R^2$ of 0.02 for ALP model). The predictive value of Mito PRS was clearly higher than the general PRS model which showed an $R^2$ of 0.04. As a sanity check step for the specificity of SNP sets contributing to PD risk, we used a PD-irrelevant pathway, i.e. myogenesis, as a negative control and a PD-relevant pathway, i.e. inflammation, as a positive control and constructed PRS models for each to assess their predictive behaviour. As expected, the myogenesis PRS model exhibited very low predictive power with an $R^2$ of 0.0045, while the inflammation PRS model had significant predictive power with an $R^2$ of 0.027 which was similar to ALP model. That the model fit of negative control PRS model was not zero could be due to the fact that the SNP weights used were still from PD GWAS and therefore not expected to be null. Furthermore, PRS model based on SNPs associated with another brain disorder, ADHD, did not show significant predictive power when applied in our target population ($p$-value $>0.05$). This supports our hypothesis regarding the mechanistic contribution of our selected pathways in PD and rules out the possibility that any group of SNPs, including those associated with other brain disorders, can be predictive of PD. Our results in terms of the prediction of PD status may provide further evidence on the mechanistic contribution of these important PD pathways to overall disease risk. This was in line with another study looked into the PD status prediction by pathway PRS (Bandres-Ciga et al., 2020).

In terms of predicting PD clinical outcomes, while a number of studies found that a higher PD PRS were significantly associated with earlier AAO (Escott-Price et al., 2015; Ibanez et al., 2017; Han et al., 2021), our results for pathway PRS were contrary to this expectation. Nevertheless, our results of mitochondrial PRS and AAO prediction replicated the work of Billingsley et al. (2019) in which they found that higher mitochondrial PRS significantly correlated with later AAO ($P$-value $<0.05$). This finding is somewhat consistent with the results of the recent and largest AAO-PD GWAS, that not all well-established risk loci are associated with AAO and suggests a likely different mechanism for PD and AAO causation. As demonstrated by Blauwendraat et al. (Blauwendraat et al., 2019) in their GWAS study of AAO in PD, the heritability of AAO in PD is approximately 11 %, whereas the heritability of PD risk is approximately 27 %, suggesting that AAO in PD is likely to be determined less by genes and more by a complex interaction with environmental factors. Therefore, prediction of AAO in PD by risk variants which are based on risk GWAS data are likely to be less informative. In terms of other clinical outcomes, such as cognitive and motor scores, a study by Paul and colleagues reported that PRS were associated with significantly faster cognitive decline, as was measured by change in the Mini-Mental State Examination (MMSE). PRS in Paul's study were also associated with faster motor decline as measured by time to Hoehn & Yahr Stage 3 and UPDRS III (Paul et al., 2018). In our study, we did not find any significant correlation between the two PRS models and MoCA and UPDRS III. One possible explanation for the lack of conclusive

prediction of PD clinical outcomes by pathway PRS in this study could be the limited number of cases with available clinical data. In addition, it is important to keep in mind that polygenic scores can only explain a small part of the actual variation in individuals. This is because only SNPs are included in the PRS construction and the rest of the genetic factors, e.g., non-SNP variants, have not yet been captured. Therefore, it would not be surprising if the current PRS, as SNP-driven entities, lack sufficient power to accurately predict the heterogeneous clinical outcomes of PD in a consistent manner. Nevertheless, it can be assumed that Pathway PRS in combination with other informative data, such as biomarkers, can contribute even better to the prediction of the clinical outcome of PD.

Regarding the use of pathway PRS in the clinic, they appear to be useful potential tools for PD patients with significant impairment in a particular pathway, i.e., high-risk individuals identified by a particular pathway PRS model and for whom there are already proven therapeutic options. As an interesting example, it has been shown that the use of certain neuroprotective therapies such as coenzyme Q10 and a number of oral medication being assessed in phase II of clinical trials can potentially improve the mitochondrial function (Day and Mullin, 2021). Similarly, it turns out that aberrant autophagy and lysosomal function can be boosted by pharmacological agents such as rapamycin (Moors et al., 2017).

Although studies of pathway PRS in PD need to be performed in much larger cohorts with a much larger number of patients with available clinical outcomes, identifying high-risk patients at significantly increased risk for particular pathway(s) and treating them with approved drugs may be a promising approach for personal medicine. In conclusion, despite all limitations, pathway PRS could play a role in biologically relevant risk stratification of patients and identification of high-risk individuals as they rely on robust results from powerful GWA studies.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mcn.2022.103751.

## CRediT authorship contribution statement

Computation, manuscript writing (MD), Review and editing (HL), Technical consultation (VB), Genotype imputation and quality control (AAKS), Statistical analysis (CS), Conceptualization, review, and editing (TG).

## Declaration of competing interest

There is no financial or non-financial conflict of interests to declare by the authors.

## References

Antony, P.M., Diederich, N.J., Krüger, R., Balling, R., 2013. The hallmarks of Parkinson's disease. FEBS J. 280 (23), 5981–5993.
Bandres-Ciga, S., Saez-Atienzar, S., Kim, J., Makarious, M., Faghri, F., Diez-Fairen, M., Iwaki, H., Leonard, H., Botia, J., Ryten, M., 2020. Large-scale pathway specific

polygenic risk and transcriptomic community network analysis identifies novel functional pathways in Parkinson disease. Acta Neuropathol. 140 (3), 341–358.

Billingsley, K.J., Barbosa, I.A., Bandrés-Ciga, S., Quinn, J.P., Bubb, V.J., Deshpande, C., Botia, J.A., Reynolds, R.H., Zhang, D., Simpson, M.A., 2019. Mitochondria function associated genes contribute to Parkinson's disease risk and later age at onset. npj Parkinson's Dis. 5 (1), 1–9.

Blauwendraat, C., Faghri, F., Pihlstrom, L., Geiger, J.T., Elbaz, A., Lesage, S., Corvol, J.-C., May, P., Nicolas, A., Abramzon, Y., 2017. NeuroChip, an updated version of the NeuroX genotyping platform to rapidly screen for variants associated with neurological diseases. Neurobiol. Aging 57, 247. e213–247. e249.

Blauwendraat, C., Heilbron, K., Vallerga, C.L., Bandres-Ciga, S., Von Coelln, R., Pihlstrøm, L., Simón-Sánchez, J., Schulte, C., Sharma, M., Krohn, L., 2019. Parkinson's disease age at onset genome-wide association study: defining heritability, genetic loci, and α-synuclein mechanisms. Mov. Disord. 34 (6), 866–875.

Calvo, S.E., Clauser, K.R., Mootha, V.K., 2016. MitoCarta2. 0: an updated inventory of mammalian mitochondrial proteins. Nucleic Acids Res. 44 (D1), D1251–D1257.

Chang, D., Nalls, M.A., Hallgrímsdóttir, I.B., Hunkapiller, J., Van Der Brug, M., Cai, F., Kerchner, G.A., Ayalon, G., Bingol, B., Sheng, M., 2017. A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. Nat. Genet. 49 (10), 1511–1516.

Choi, S.W., O'Reilly, P.F., 2019. PRSice-2: Polygenic Risk Score software for biobank-scale data. Gigascience 8 (7), giz082.

Choi, S.W., Mak, T.S.-H., O'Reilly, P.F., 2020. Tutorial: a guide to performing polygenic risk score analyses. Nat. Protoc. 15 (9), 2759–2772.

Day, J.O., Mullin, S., 2021. The genetics of Parkinson's disease and implications for clinical practice. Genes 12 (7), 1006.

Demontis, D., Walters, R.K., Martin, J., Mattheisen, M., Als, T.D., Agerbo, E., Baldursson, G., Belliveau, R., Bybjerg-Grauholm, J., Bækvad-Hansen, M., 2019. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. Nat. Genet. 51 (1), 63–75.

Dudbridge, F., 2013. Power and predictive accuracy of polygenic risk scores. PLoS Genet. 9 (3), e1003348.

Duncan, L., Shen, H., Gelaye, B., Meijsen, J., Ressler, K., Feldman, M., Peterson, R., Domingue, B., 2019. Analysis of polygenic risk score usage and performance in diverse human populations. Nat. Commun. 10 (1), 1–9.

Escott-Price, V., , I. P. s. D. G. Consortium, Nalls, M.A., Morris, H.R., Lubbe, S., Brice, A., Gasser, T., Heutink, P., Wood, N.W., Hardy, J., 2015. Polygenic risk of P arkinson disease is correlated with disease age at onset. Ann. Neurol. 77 (4), 582–591.

Finkbeiner, S., 2020. The autophagy lysosomal pathway and neurodegeneration. Cold Spring Harb. Perspect. Biol. 12 (3), a033993.

Fung, H.-C., Scholz, S., Matarin, M., Simón-Sánchez, J., Hernandez, D., Britton, A., Gibbs, J.R., Langefeld, C., Stiegert, M.L., Schymick, J., 2006. Genome-wide genotyping in Parkinson's disease and neurologically normal controls: first stage analysis and public release of data. Lancet Neurol. 5 (11), 911–916.

Gasser, T., 2007. Update on the genetics of Parkinson's disease. Mov.Disord. 22 (S17), S343–S350.

Han, Y., Teeple, E., Shankara, S., Sadeghi, M., Zhu, C., Liu, D., Wang, C., Frau, F., Klinger, K., Madden, S., 2021. Genome-wide Polygenic Risk Score Identifies Individuals at Elevated Parkinson's Disease Risk, 4286. AAN Enterprises.

Hou, X., Watzlawik, J.O., Fiesel, F.C., Springer, W., 2020. Autophagy in Parkinson's disease. J. Mol. Biol. 432 (8), 2651–2672.

Huang, J., Ling, C.X., 2005. Using AUC and accuracy in evaluating learning algorithms. IEEE Trans. Knowl. Data Eng. 17 (3), 299–310.

Ibanez, L., Dube, U., Saef, B., Budde, J., Black, K., Medvedeva, A., Del-Aguila, J.L., Davis, A.A., Perlmutter, J.S., Harari, O., 2017. Parkinson disease polygenic risk score is associated with Parkinson disease status and age at onset but not with alpha-synuclein cerebrospinal fluid levels. BMC Neurol. 17 (1), 1–9.

Lewis, C.M., Vassos, E., 2020. Polygenic risk scores: from research tools to clinical instruments. Genome Med. 12, 1–11.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., Tamayo, P., 2015. The molecular signatures database hallmark gene set collection. Cell Syst. 1 (6), 417–425.

Loza, M.J., McCall, C.E., Li, L., Isaacs, W.B., Xu, J., Chang, B.-L., 2007. Assembly of inflammation-related genes for pathway-focused genetic analysis. PloS one 2 (10), e1035.

Moors, T.E., Hoozemans, J.J., Ingrassia, A., Beccari, T., Parnetti, L., Chartier-Harlin, M.-C., Van De Berg, W.D., 2017. Therapeutic potential of autophagy-enhancing agents in Parkinson's disease. Mol. Neurodegener. 12 (1), 1–18.

Nalls, M.A., Blauwendraat, C., Vallerga, C.L., Heilbron, K., Bandres-Ciga, S., Chang, D., Tan, M., Kia, D.A., Noyce, A.J., Xue, A., 2019. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. Lancet Neurol. 18 (12), 1091–1102.

Pan, T., Kondo, S., Le, W., Jankovic, J., 2008. The role of autophagy-lysosome pathway in neurodegeneration associated with Parkinson's disease. Brain 131 (8), 1969–1978.

Park, J.-S., Davis, R.L., Sue, C.M., 2018. Mitochondrial dysfunction in Parkinson's disease: new mechanistic insights and therapeutic perspectives. Curr.Neurol. Neurosci.Rep. 18 (5), 21.

Paul, K.C., Schulz, J., Bronstein, J.M., Lill, C.M., Ritz, B.R., 2018. Association of polygenic risk score with cognitive decline and motor progression in Parkinson disease. JAMA Neurol. 75 (3), 360–366.

Purcell, S., Chang, C. PLINK2 (v1. 90b6. 9). Available online. (Accessed 5 January 2019).

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer.