Precision Medicine for Alzheimer's Disease: Identifying Effective Therapeutic Strategies with Patient Stratification Biomarkers

K. Taylor, A. Malinowski, S. Das
PrecisionLife Ltd, Oxford, United Kingdom

Scan to view online:

Introduction

PrecisionLife is a pioneering techbio company, with a unique approach that finds more signal in complex disease patient data than standard methods. Our high-resolution patient stratification identifies subgroups of patients with similar disease drivers and treatment responses, to make precision medicine possible in chronic diseases.

Alzheimer's disease (AD), like other complex diseases, is characterized by a high degree of heterogeneity across the patient population, reflected in a wide range of disease presentations and therapy responses. GWAS have identified several disease-associated genes, but these findings have not translated into progress in clinical trials. This likely reflects the limitations of GWAS in only identifying single variants, while the key to understanding complex diseases that are influenced by multiple genetic loci is to find combinations of variants that distinguish one patient subgroup from another.

Methods

DATASET:
PrecisionLife analysed a dataset constructed from the UK Biobank:

Alzheimer's disease Cases (n = 882):
- Alzheimer's disease diagnosis (ICD-10 code, G30.x)

Healthy Controls (n = 1,816):
- No reported neurodegenerative disorders
- No self-reported cognitive decline
- No family history of Alzheimer's disease

COMBINATORIAL ANALYSIS:
The dataset was analysed in the PrecisionLife platform to identify combinations of SNP genotypes that when observed together in a patient are strongly associated with AD.

SNP combinations that have high odds ratios, low *p*-values and high prevalence in cases are prioritized. This process undergoes 1,000 cycles of fully randomized permutations and combinations must meet a specified FDR threshold.

SNPs are scored using a Random Forest algorithm in a 5-fold cross-validation framework and prioritized based on their ability to differentiate cases and controls.

The highest scoring SNPs are then mapped to genes and clustered by the patients they co-occur in to generate a disease architecture.

Results

Figure 1. Combinatorial analysis of genomic data vs GWAS

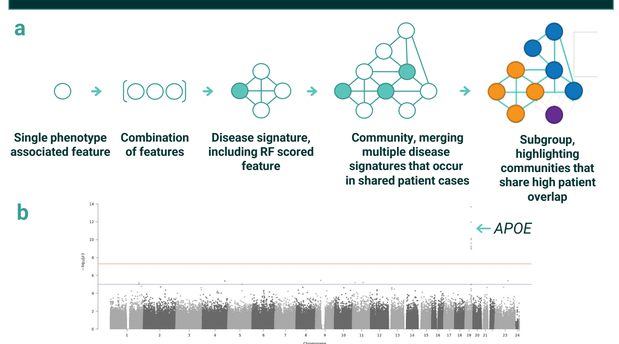

Figure 1. (a) Conceptual representation of features, combinations, disease signatures and communities used to build up the disease architecture in the PrecisionLife combinatorial methodology. (b) Manhattan plot of genome-wide *p*-values of association for the AD UK Biobank cohort. The dashed line represents the genome-wide significance threshold at p<5e-08.

GWAS	Combinatorial Analysis
Single SNP associations must be significant across large groups of patients	Specific combinations of variants associated with each patient subgroup serve as a genetic stratification biomarker
Limited insights unless disease is likely to be caused by a small number of rare variants with large effect sizes (often in gene coding regions affecting protein 3D structure)	Patient subgroups with different causes of disease or even incorrect diagnoses can be distinguished (stratified) by different mechanistic etiology
Does not account for the effects of interactions between SNPs, genes and metabolic networks	Captures epistatic and non-linear additive effects of all interactions between SNPs, genes, environmental factors and metabolic networks

Table 1. Key results from PrecisionLife's AD study

Validated disease signatures (SNP combinations)	4,887
Significant RF-scored SNPs associated with AD	267
Significant genes associated with AD	113 (inc. *APOE* and *APOC1*)
Genes targeted by at least one drug in clinical development*	32
Patient communities identified	13
Patient subgroups (merged communities) identified	6

*source: DrugBank and ChEMBL

Table 1. Summary of the results of PrecisionLife's combinatorial analysis of the UK Biobank Alzheimer's disease genomic dataset.

Figure 2. Alzheimer's disease architecture reveals 13 stratified genetic communities of patients

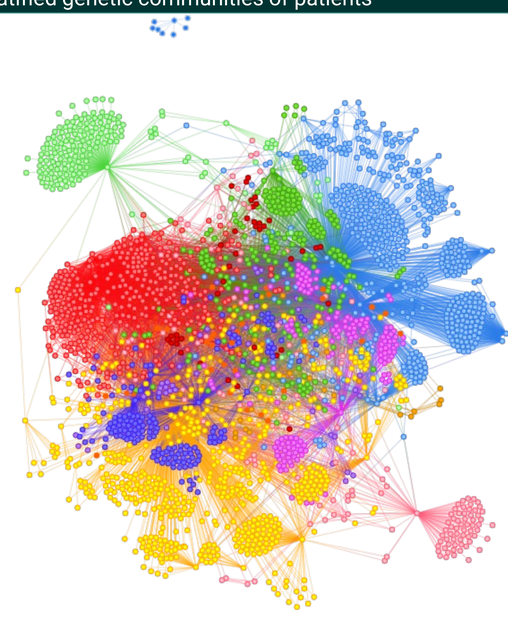

Figure 2. Disease architecture diagram demonstrating the 13 communities of SNPs comprising the structure of the Alzheimer's disease patient subpopulations generated by the PrecisionLife platform.
Each circle represents a disease-associated SNP genotype, edges represent co-association in patients, and each color a distinct community of SNPs.

Figure 3. Patient overlap analysis merges communities into six major genetic subgroups with distinct mechanisms

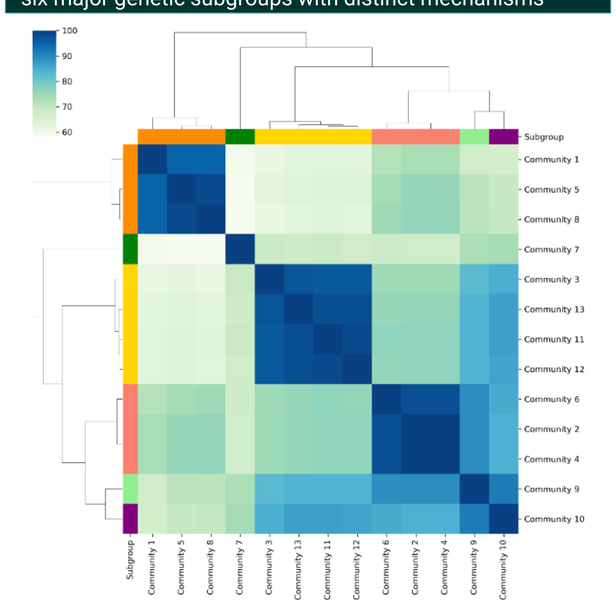

Figure 3. Clustered heatmap showing the overlap of AD patients associated with 13 communities. Each border color represents a patient subgroup (n = 6)

Figure 4. Six patient subgroups are associated with distinct biological pathways and pathological mechanisms

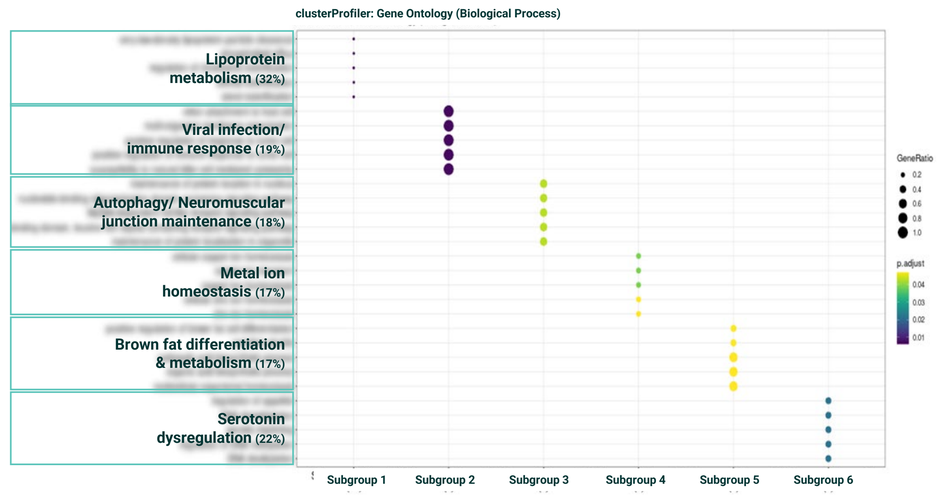

Figure 4. Pathway enrichment plot for the genes found in the communities associated with different patient subgroups. Gene ratio represents the ratio of genes found in the pathway compared to the genes associated with a community and *p*.adjust represents the *p*-value adjusted for multiple testing. The dots in the plot are color-coded based on their corresponding *p*.adjust values. The percentage of cases represented by each subgroup is displayed in brackets, patients may belong to multiple subgroups.

Figure 5. Pathways identified in PrecisionLife's AD study are also significant in other CNS diseases

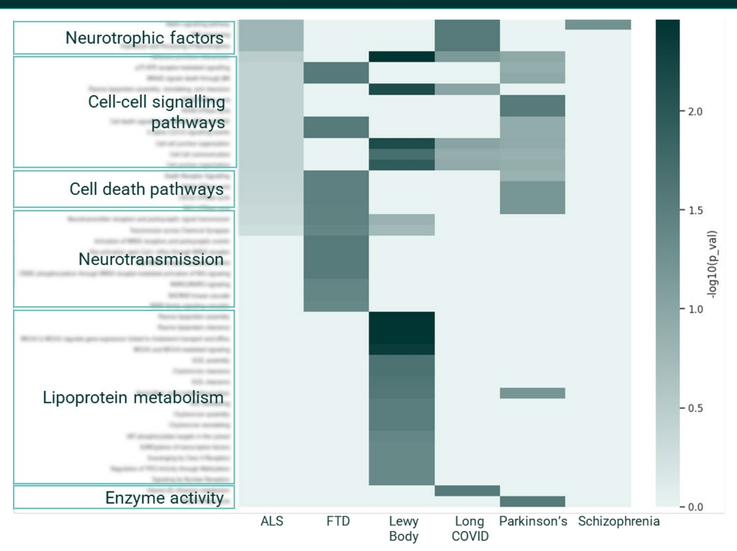

Figure 5. Pathway enrichment for genes identified in AD and other CNS diseases analysed by PrecisionLife. Enrichment analysis was performed using the GProfiler tool to determine pathways significantly enriched in at least one indication (p<0.05, *p*-value correction for multiple testing using 'Benjamini-Hochberg'). Scores shown are -log10(*p*-value).

Discussion

PATIENT STRATIFICATION
Clustering these combinations, based upon the degree of patient overlap, revealed six distinct subgroups of patients. Each patient subgroup reflected a specific biological function:
1. Lipid metabolism
2. Neuroinflammation
3. Autophagy
4. Metal ion homeostasis
5. Metabolic dysfunction
6. Serotonin signaling

NOVEL TARGETS
Our analysis identified combinations of genetic variants which mapped to 113 genes that are significantly associated with AD development. Using further analysis, we shortlisted 26 novel targets to be prioritized based on MoA hypothesis, population prevalence, localization and druggability. Selected targets will be validated in AD-relevant human iPSC-derived neurons using genetic and/or pharmacological approaches. Positively validated targets will be then tested *in vivo* disease models (e.g., mouse/zebrafish).

DRUG REPURPOSING
Among the genes identified within each AD subgroup, 32 are targeted by drugs in clinical development in other indications. We have developed a pipeline to systematically evaluate the potential of repurposing these to accelerate implementation of safe and effective therapies for AD patients.

CROSS DISEASE ANALYSIS
Several of the pathways identified are also found in patient subgroups in other neurological diseases, including FTD, schizophrenia and long COVID, indicating shared genetic aetiologies and underlying dysregulated processes common between them.

Conclusion

The results demonstrate that the PrecisionLife combinatorial analysis is uniquely able to stratify heterogenous patient populations with complex disease pathologies. We can use these insights to identify more effective therapeutic strategies and accompanying biomarker sets to match them to the patient subgroups that are most likely to demonstrate benefit in downstream clinical trials.

Acknowledgements
Research described in this study has been conducted using data from the UK Biobank Resource (application number 44288).

precisionlife
improving health | for everyone

For more information, visit:
www.precisionlife.com