Characterizing the Genetic and Biological Differences between Endometriosis and Adenomyosis using the UK Biobank

Krystyna Taylor[1], Karolina Chocian[1], Sayoni Das[1], Jason Sardell[1], Amy Rochlin[2], Steve Gardner[1]
[1]PrecisionLife, Oxford, United Kingdom, [2]Complex Disorders Alliance, Stamford, CT, USA

Introduction

PrecisionLife has developed a unique combinatorial approach to analyzing large scale genomic and other patient data. This captures the non-linear effects of interactions between multiple genes and exogenous (e.g., clinical, epidemiological, transcriptomic etc.) factors. The inclusion of non-linear interactions enables discovery of deeper and more reproducible insights than Genome-Wide Association Studies (GWAS) alone.

We find more significant SNPs in complex disease patient data than GWAS, evaluate their causality, explain more disease variance, and our results translate between populations with different ancestries better than GWAS or polygenic risk scores[1,2].

Our mechanistic patient stratification correlates disease risk signatures with specific subgroups of patients who share similar disease drivers and treatment responses, to make precision medicine possible in over 60 complex chronic diseases. We work with key opinion leaders, disease charities and patients to find better treatment options for patients with unmet medical needs, including endometriosis and other chronic women's health disorders.

Endometriosis & Adenomyosis

Endometriosis and adenomyosis are chronic gynecological conditions that significantly impact the quality of life of hundreds of millions of women. Adenomyosis and endometriosis produce similar debilitating symptoms, including dysmenorrhea and chronic pelvic pain. However, they differ in the underlying pathology as in endometriosis, endometrial tissue growth occurs outside the uterus, while in adenomyosis, it invades the uterine walls[3].

Patients often face years-long diagnostic delays, inadequate symptom management, and insufficient research funding has led to persistent unmet needs in awareness, diagnosis, and comprehensive care.

We aimed to identify genetic signatures in both diseases and evaluate the genetic and biological differences between them in order to further our understanding of the overlapping and unique disease mechanisms driving each condition.

Methods

UK BIOBANK[4] DATASETS

This study applied combinatorial analytics to genotype data for 2 case cohorts, using ICD-10 codes to designate adenomyosis or endometriosis case status.

- Adenomyosis (ICD-10 code, N80.0) – n = 2,024
- Endometriosis (ICD-10 code, N80.1-9) – n = 4,493

Each case group was compared against a unique healthy female control population without chronic gynecological disease.

COMBINATORIAL ANALYSIS:

Each dataset was analyzed in the PrecisionLife platform to identify combinations of SNP (single nucleotide polymorphisms) genotypes that were significantly associated with the case population (i.e., endometriosis or adenomyosis) (Figure 1).

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.

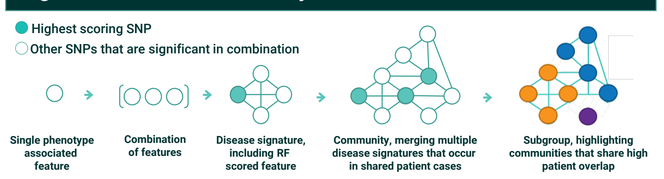
SNPs are scored using a Random Forest (RF) algorithm in a 5-fold cross validation framework and prioritized based on their ability to differentiate cases and controls.

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

Reproduction of disease signatures (Figure 2) was performed in All of Us[5] case-control cohorts.

Results

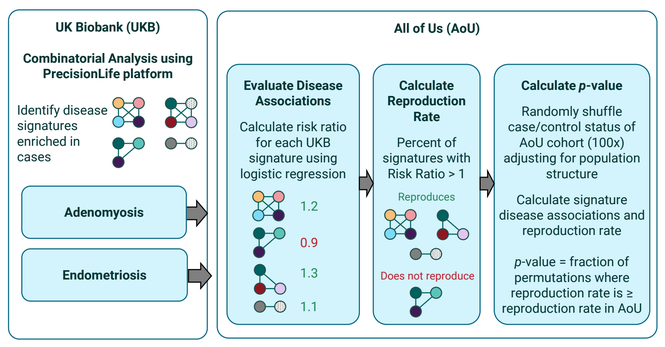
Figure 1. Combinatorial Analysis of Genomic Data vs GWAS



Conceptual representation of features, combinations, disease signatures and communities used to build up the PrecisionLife disease architecture

GWAS	Combinatorial Analysis
Single SNP associations must be significant across large groups of patients, and they must be accurately diagnosed	Specific co-associated combinations of variants serve as a genetic stratification biomarker for each patient subgroup
Limited insights unless disease is caused by a small number of rare variants with large effect sizes (often in coding regions affecting protein 3D structure)	Patient subgroups with different causes of disease or even incorrect diagnoses can be distinguished (stratified) by different mechanistic aetiology
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

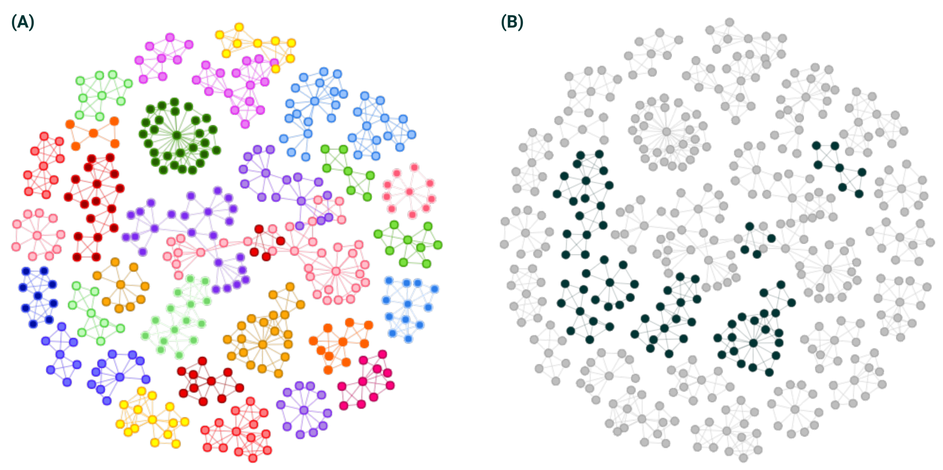
Figure 2. Methodology Pipeline using UKB and AoU Cohorts



Study flow and overview reproducing disease signatures identified in the original combinatorial analysis of the endometriosis and adenomyosis cohorts derived from the UKB in the disjoint and more ancestrally diverse All of Us cohort.

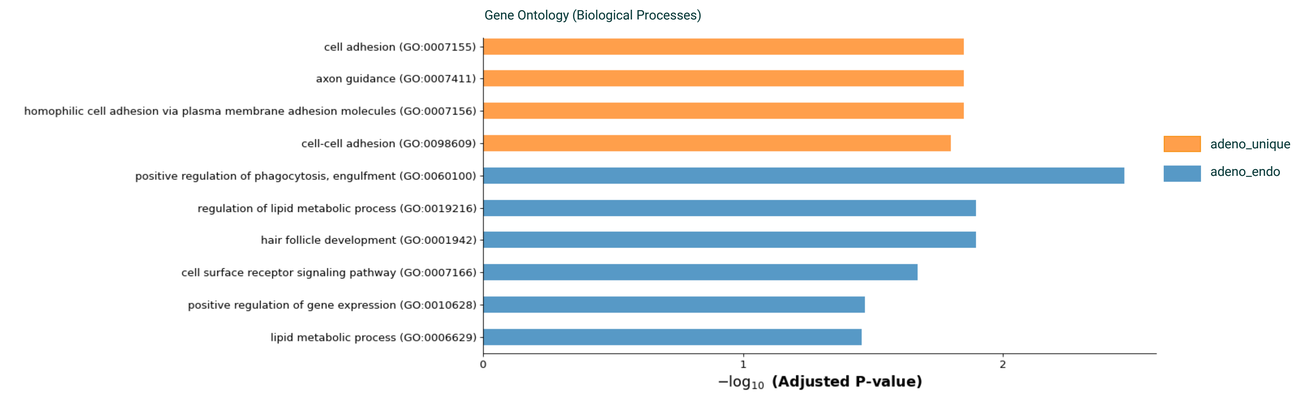
Figure 3. Combinatorial Analysis of Adenomyosis Cohort in the UK Biobank

Total disease signatures	171
Unique SNPs (RF scored)	46
Unique genes (RF scored)	27
Genes also identified in PL endometriosis study	9 (33%)
Genes with literature evidence in endometriosis	7 (26%)



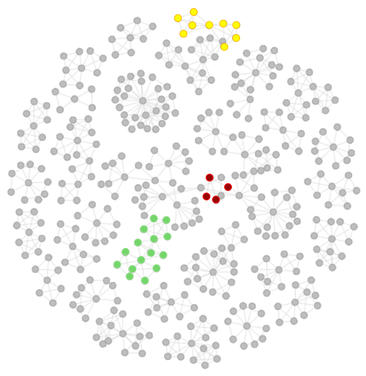
Combinatorial analysis and mechanistic patient stratification results, shown as a disease architecture, for the UK Biobank adenomyosis dataset.
Circles are SNPs, and lines indicate co-association in over 20% of patients in the study. (A) Disease architecture is coloured by patient subgroups, formed by clustering disease signatures by the patients they co-occur in. (B) Disease architecture demonstrating the genetic overlap between PL's adenomyosis and endometriosis studies. Patient subgroups containing genes also found in PL's endometriosis study are coloured black, patient subgroups containing genes unique to adenomyosis are grey.

Figure 4. Pathway Enrichment Analysis Shows Biological Differences between Disease Genes



Comparative pathway enrichment plot for genes associated with adenomyosis only ('adeno_unique') or found to also be significant in prior PL endometriosis studies ('adeno_endo'), using Gene Ontology (GO) biological processes. 'Adjusted *p*-value' represents the *p*-value adjusted for multiple testing.

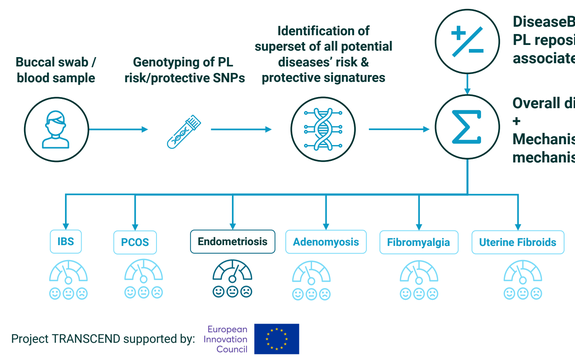
Figure 5. Examples from Gene-Subgroup Analysis of Adenomyosis Targets



SUBGROUP	CRITICAL GENE	LITERATURE EVIDENCE	PREVALENCE IN ADENOMYOSIS	PREVALENCE IN ENDOMETRIOSIS	REPRODUCTION IN All of Us
22	*PPARG*	• Promoter region hypermethylated in adenomyosis, resulting in decreased expression in adenomyotic lesions[6]	5%	N/A	**Disease signature has been reproduced in AoU (RR>1)**
31	*DSCAM*	• Decreased expression of *DSCAM* in endometriotic lesions in context of netrin-1 associated pain[7]	< 5%	20%	**Disease signature has been reproduced in AoU (RR>1)**
19	*MPPED2*	• GWAS association - endometriosis[8] • Differentially expressed in endometriosis tissue[9]	6%	22%	**Disease signature has been reproduced in AoU (RR>1)**

Table demonstrating examples of gene targets identified in 3 different patient subgroups (highlighted on the disease architecture) defined by combinations of SNPs that are significantly associated with adenomyosis. The table contains the (minimum) prevalence of the disease signature in PL's adenomyosis and endometriosis studies (if found) . Validation of disease signatures in the All of Us (AoU) dataset is defined as signatures with a risk ratio (RR) > 1.

Figure 6. Project TRANSCEND – Clinical Studies to Validate Proofs of Mechanisms



Project TRANSCEND supported by: European Innovation Council

PrecisionLife's Vision of Precision and Preventative Medicine

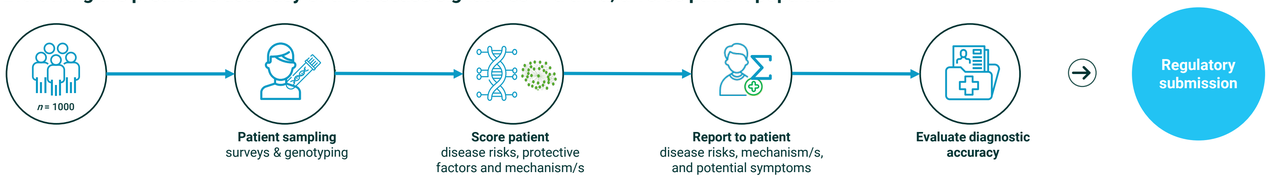
Highly predictive disease risk signatures enable non-invasive, accurate differential triage tool for endometriosis and other cause of pelvic pain

Earlier, more accurate differential triage of endometriosis against range of conflicting diagnoses e.g. adenomyosis, PCOS, IBD etc.

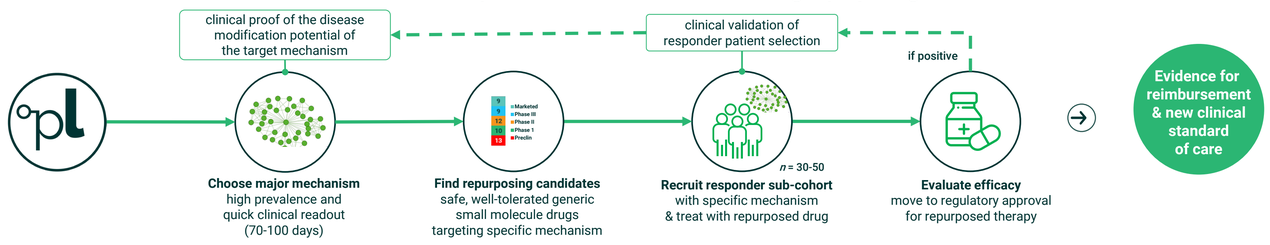
Drug repurposing candidates specific to mechanistic sub-groups for personalised treatments

PL's highly predictive risk signatures enable a new type of rapid, non-invasive disease risk test for complex chronic diseases. A test to count the number of risk signatures in a patient's makeup can be run from a single saliva sample on a standard genotyping array. These can report both the patient's disease risk and their likely mechanistic cause of disease (and therefore drug response).

Evaluating the predictive accuracy of the disease signatures in a third, diverse patient population:



Sub-selection of cohort/s based on endomethriosis specific mechanisms for which drug repurposing candidates have been identified:



Discussion

Combinatorial analysis of the adenomyosis case-control dataset revealed 171 disease-associated SNP combinations ('disease risk signatures'), mapping to 27 unique genes (Figure 3). In contrast, we identified 182 disease signatures in our combinatorial analysis of the UKB endometriosis cohort that were confirmed significant in AoU.

Nine of these genes appeared in both adenomyosis and endometriosis analyses, though in different combinations of SNP signatures. Among them, three genes, including *MPPED2* and *PPARG,* have also been associated with endometriosis in prior GWAS studies. In genes that were found in both adenomyosis and endometriosis studies, we find higher case prevalence in the UK Biobank endometriosis population as shown in Figure 5.

Genes uniquely linked to adenomyosis in this study have not been identified in previous endometriosis GWAS or non-genomic studies, suggesting novel pathways. Many of these genes are involved in uterine function, particularly through estrogen signaling, and have connections to endometrial cancer, pregnancy, and fertility.

A comparative pathway enrichment analysis suggests that there are biological differences between the genes that are unique to adenomyosis vs those common between the two (Figure 4). Genes found only in adenomyosis were enriched for pathways involved in cell-cell adhesion, whereas genes that had also been found in PL's endometriosis study were enriched for processes such as lipid metabolism and regulation of phagocytosis.

PrecisionLife has been awarded a €2.5m non-dilutive grant by the European Innovation Council (EIC) for its TRANSCEND project, to more rapidly and accurately triage and treat patients with endometriosis. As adenomyosis and endometriosis share similar symptoms, and may be co-associated in some women, it is crucial to further understand the genetic and mechanistic overlap between the two disorders. Reproduction of both adenomyosis and endometriosis disease signatures in the AoU and other datasets is critical to provide further *in silico* validation for the significance and predictive value of these signatures in defining each disease.

Conclusion

The results demonstrate that the PrecisionLife combinatorial analysis platform is uniquely able to stratify heterogenous patient populations with complex disease pathologies. We can use these insights to identify more effective diagnostic strategies and accompanying therapies.

Adenomyosis remains challenging to diagnose and treat. This research provides valuable insights into the genetic risk factors and pathophysiological mechanisms underlying the disease, potentially improving diagnostic and therapeutic approaches in adenomyosis.

References

1. Taylor, K., Pearson, M., Das, S., Sardell, J., Chocian, K., Gardner, S. Genetic risk factors for severe and fatigue dominant long COVID and commonalities with ME/CFS identified by combinatorial analysis. *J Transl Med* **21**, 775 (Nov 2023). https://doi.org/10.1186/s12967-023-04588-4
2. Sardell, J., et al., Reproducibility of Genetic Risk Factors Identified for Long COVID across US and UK Patient Cohorts (in press) *J Transl Med* https://www.researchsquare.com/article/rs-5992226/v1
3. Donnez J, Stratopoulou CA, Dolmans MM. Endometriosis and adenomyosis: Similarities and differences. *Best Pract Res Clin Obstet Gynaecol.* 2024;**92**:102432. https://doi.org/10.21037/atm-20-2161
4. Sudlow C, Gallacher J, Allen N, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015;**12**(3):e1001779. Published 2015 Mar 31. https://doi.org/10.1371/journal.pmed.1001779
5. All of Us Research Program Investigators, Denny, J. C., Rutter, J. L., Goldstein, D. B., Philippakis, A., Smoller, J. W., Jenkins, G., & Dishman, E. (2019). The "All of Us" Research Program. *New England Journal of Medicine*, **381(7)**, 668–676. https://doi.org/10.1056/NEJMsr1809937
6. Fan, J., Liu, X. & Guo, SW. Hypermethylation of Klotho and Peroxisome Proliferator-Activated Receptor γ Concomitant with Overexpression of DNA Methyltransferase 1 in Adenomyosis. *Reprod. Sci.* **32**, 668–683 (2025). https://doi.org/10.1007/s43032-024-01599-4
7. Ding, S., Guo, X., Zhu, L., Wang, J., Li, T., Yu, Q., & Zhang, X. (2021). Macrophage-derived netrin-1 contributes to endometriosis-associated pain. *Annals of Translational Medicine*, **9**(1), 29. https://doi.org/10.21037/atm-20-2161
8. OpenTargets (2024). https://platform.opentargets.org/evidence/ENSG00000066382
9. Mirza Z, Abdel-Dayem UA. Uncovering Potential Roles of Differentially Expressed Genes, Upstream Regulators, and Canonical Pathways in Endometriosis Using an In Silico Genomics Approach. *Diagnostics* (Basel). 2020;**10**(6):416. Published 2020 Jun 19. https://doi.org/10.3390/diagnostics10060416

Acknowledgements

Research described in this study was conducted using data from the UK Biobank Resource (application 44288) and All of Us Research Program accessed in collaboration with the Complex Disorders Alliance (CODA).

Thanks to all patients who gifted their data and the whole PrecisionLife team for their input and support.

precisionlife
improving health | **for everyone**

For more information, visit:
www.precisionlife.com