First Year Review of the LOCOME Project:
Progress in Understanding ME/CFS and Long COVID and Finding New Diagnostic and Therapeutic Options

Steve Gardner, PrecisionLife Ltd

Executive Summary

ME/CFS and long COVID have affected the lives of over 400 million people worldwide. The diseases are now estimated to contribute to over $1 trillion of healthcare costs and lost economic productivity[1,2] – equivalent to 1% of global GDP. But in spite of their huge personal and public health impacts, we still have no tools to accurately diagnose patients and no drugs that treat the underlying causes of the diseases. This has led to widespread misunderstanding, lack of awareness, and even denial of the diseases in the clinical and social care communities, which further harms sufferers.

In large part this is because in spite of concentrated global efforts investigating long COVID, there was no fully replicated and peer reviewed evidence for any genes associated with the disease. Understanding the biological causes of ME and long COVID is particularly difficult because of the large number of different disease symptoms and organs that are affected[3].

The UK-based LOCOME project and our US-based MELO study are changing this. We're building a deeper understanding of the different factors causing the diseases, showing where they're similar and different, and explaining why they produce such a range of symptoms in patients. We've had some very encouraging results from the first year of studying the genetics of people with these diseases:

- **We found 14 novel genes in ME/CFS (UK Biobank) and 73 in long COVID (Sano GOLD)**
- **We have replicated between 77-94% of these genes in separate patient populations including the UK DecodeME and the US All of Us cohorts.**
- **7 of the 9 genes we found originally that overlapped between our ME/CFS and long COVID studies were reproduced in the All of Us dataset (the other 2 could not be tested). There is very good evidence now of some shared pathophysiology between these disease.**
- **We've identified 9 safe generic drugs that we hope to prove may benefit patients in both diseases – like the COVID-19 RECOVERY repurposing trial but with genetic evidence.**
- **We found 9 genes we believe may act to resist the disease mechanism and protect people against the onset / progression of the diseases. These are a new class of drug targets that cannot be found using existing methodologies.**
- **We're turning these insights into a new class of rapid, non-invasive diagnostic tests based on a buccal swab that will determine a patient's disease risk, and the mechanism/s underpinning their disease.**
- **We are using this to enable differential triage across multiple post-viral syndromes and other indications with similar symptoms and help guide therapy selection.**
- **We're running clinical trials in the US, recruiting 1,000 ME and long COVID patients to clinically validate these new diagnostic & therapeutic options.**

This work is published in multiple papers referenced in the text below.

[1] Al-Aly, Z., Davis, H., McCorkell, L. et al. Long COVID science, research and policy. *Nature Medicine* 30, 2148–2164 (2024).

[2] Bateman Horn Center / Optum Estimating Prevalence, Demographics, and Costs of ME/CFS (2019) *Front. Pediatr.* 8;6:412

[3] Long-Term Long Covid (2023) https://erictopol.substack.com/p/long-term-long-covid

Background

PrecisionLife's work in these diseases started in May 2020 on severe COVID-19 with the first UK Biobank cohort[4]. **We found 68 genes associated with the disease that explain the unexpectedly wide range of symptoms shown by COVID-19 patients, and several drug repurposing candidates, one of which was shown to reduce disease severity and duration by 40-45%[5].** We showed that the disease mechanisms we highlighted could predict which patients were most likely to need oxygen support and intensive care in a study in United Healthcare's 300,000 US COVID-19 patients in March 2021[6].

In December 2022 we ran a study of two separate ME/CFS populations from UK Biobank[7]. This used our innovative new way of analyzing patients' genetic and clinical data using our AI platform (combinatorial analysis[8]), which is particularly well suited to finding deeper genetic associations than traditional analysis methods and explaining more about how complex diseases manifest in patients.

A key difference to understand about our approach is that it's based on our belief that there are many different causes of a disease across patients. They may be given the same diagnosis just because their symptoms appear similar, e.g., they all have fatigue, but that doesn't mean their diseases are the same. It's like there being many ways to get to central London – the result is the same, but the starting points and routes taken can be different. This means that drugs that benefit one patient may not work in another if their cause of disease (their disease 'mechanism') is different – just like running twice as many trains into London from Oxford wouldn't have any impact on people travelling into the city from Brighton. Identifying the specific personal mechanisms driving disease in a patient is crucial to choosing their most effective personalized treatment regimen.

What We've Found So Far - ME

In this first analysis[7], we found 199 genetic mutations (single nucleotide polymorphisms or SNPs) in 14 genes that were associated with ME/CFS in two UK Biobank populations. Some appeared in both populations – providing extra evidence that these genes are likely to be important in the disease.

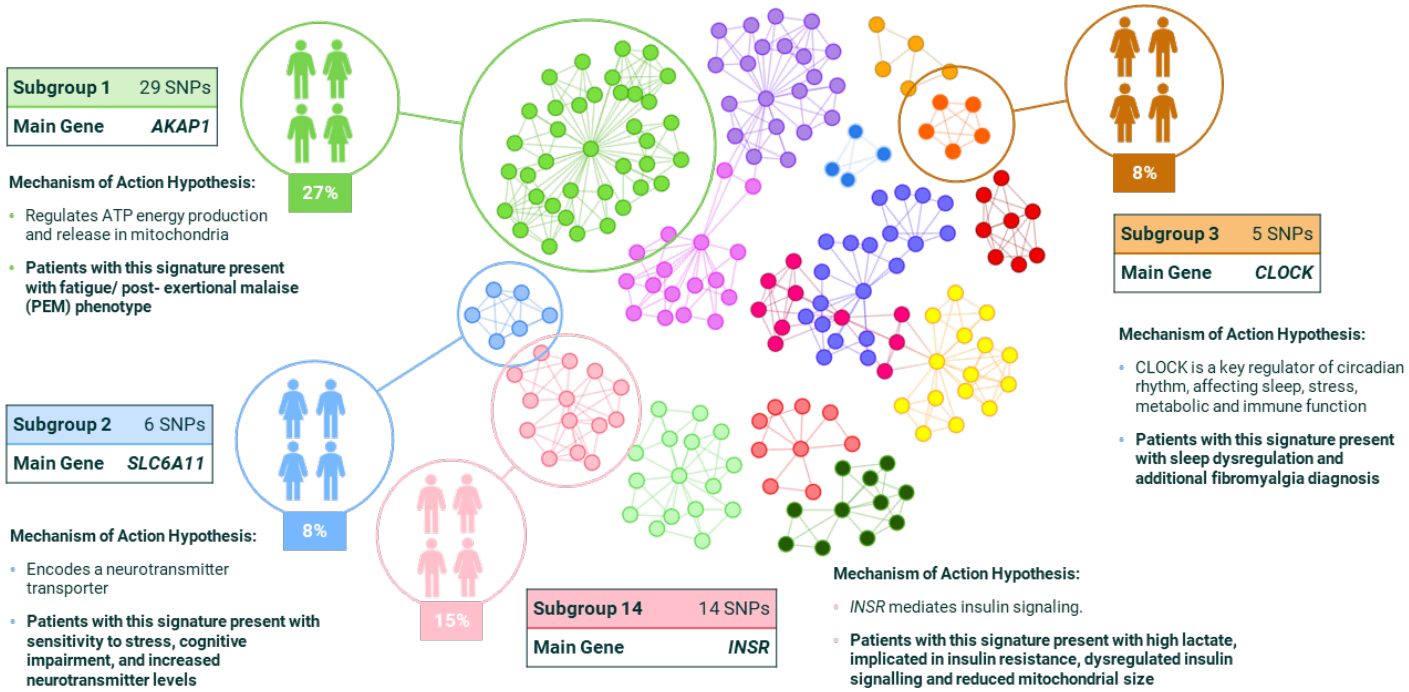

Figure 1. Patient stratification of ME/CFS, testing that patient subgroups with specific genetic disease signatures (the coloured clusters) show symptoms that are consistent with the known functions of the genes in those disease signatures. The results are highly statistically significant with odds ratios between 3.2 and 4.7 (odds ratios above 1.4 are considered very predictive).

[4] Taylor, K., et al. Analysis of Genetic Host Response Risk Factors in Severe COVID-19 Patients medRxiv 2020.06.17.20134015;

[5] Cadegiani FA, et al. Early Antiandrogen Therapy With Dutasteride Cureus. (2021) 13(2):e13047.

[6] Das, S., et al., 2021. Combinatorial analysis of phenotypic and clinical risk factors associated with hospitalized COVID-19 patients. *Front Digit Health*. 3, 660809.

[7] Das, S., et al., 2022. Genetic risk factors for ME/CFS identified using combinatorial analysis. *J. Trans Med*, 20(1), p.598

[8] Gardner, S., 2021. Combinatorial analytics. Artificial Intelligence in the Life Sciences, 1, p.100003.

To check that these results made sense, we looked at the subgroups of patients who had mutations in these genes to see if their symptoms were consistent with what we know about the role (mechanism of action) of that gene in the body. A good example is the gene *AKAP1*, which is involved in recharging the cell's energy supply after it's been used up by exercise (like recharging a battery). We checked if those patients who had mutations in this gene (the green subgroup top left in Figure 1) were in fact reporting fatigue and post-exertional malaise (PEM) as major symptoms of their disease. It was reassuring that these patients did indeed report PEM as a symptom in significant numbers.

Building on these encouraging initial findings, we applied for and won funding from Innovate UK for the LOCOME project, which kicked off in December 2023. Led by PrecisionLife, with Prof. Chris Ponting's group at University of Edinburgh, and Action for ME as partners, this project was designed to confirm the earlier UK Biobank findings using DecodeME (https://www.decodeme.org.uk), the world's largest ME/CFS study, and other global data sources. We've since replicated many of the findings of the original study in the first batches of DecodeME data, and are working with the DecodeME team to complete and publish the results of this work.

What We've Found So Far – Long COVID

Our work in long COVID started in mid-2023 when we accessed Sano Genetics' GOLD long COVID patient dataset. We looked at two sets of <500 long COVID patients who reported fatigue and/or very severe forms of the disease. **We identified over 5,000 SNPs and 73 genes associated with long COVID[9].** This is massively more than has been found by any other study to date, even those that have analyzed over 50,000 patients[10].

Several of those genes aligned with neurological, cardiovascular and metabolic diseases as expected, and **9 of them were also found in the original 14 genes from the ME/CFS analysis**. These included for example genes involved in circadian rhythm regulation and insulin regulation, which are closely associated with key ME/CFS symptoms (subgroup 14 in pink and subgroup 3 in orange in Figure 1).

Figure 2. Numbers of genes identified (and reproduced) in ME/CFS and long COVID, and the overlap between the diseases.

We checked the 73 long COVID genes we found originally in Sano GOLD, to see if they were also significantly associated with long COVID in the All of Us cohort from the US. This is important, as it's a much more diverse population – UK Biobank, Sano GOLD and DecodeME are all well over 90% White European ancestry, but All of Us is only 55% White European with other genetic ancestries, including African, Asian and Latino participants in greater numbers. It's very important to have these ancestries represented as it widens the relevance and credibility of the results and enhances health equity.

We were very encouraged by our recent results, which show that **we can reproduce 77%-83% of the SNPs and 83%-93% of the long COVID genes that we originally found in Sano GOLD in the All of Us dataset[11].** These numbers do go down when considering non-White ancestries – to 51%-64% in Black/African-American and Hispanic/Latino populations, but this is still a much higher proportion of findings reproducing than expected from such a biased starting dataset. We're obviously aiming to improve this number, but it still strongly suggests that the results will have widespread benefit across patients with different ancestries. **Reproducing these results is a very positive indication that the genes that we have found are indeed associated with increasing long COVID disease risk**. Of the 9 long COVID genes that overlapped with ME/CFS, 7 were also seen in the All of Us long COVID cohort.

We predicted 13 targets for drug repurposing in our original long COVID study, and 11 of these were reproduced in the All of US cohort. This includes all the targets that are in common between ME/CFS and long COVID, giving extra priority to drugs targeting these genes in our repurposing clinical trials.

---

[9] K. Taylor, et al., 2023. Genetic risk factors for long COVID and commonalities with ME/CFS identified by combinatorial analysis. *J Transl Med*. 21, 775.

[10] Multi-ancestry GWAS of Long COVID Chaudhary, NS, Weldon, CH et al. medRxiv 2024.10.07.24315052;

[11] Sardell, J., et al., 2025 Reproducibility of genetic risk factors identified for long COVID using combinatorial analysis across US and UK patient cohorts Study in press – https://www.medrxiv.org/content/10.1101/2025.02.04.25320937v1

How Can We Use These Results?

This is very encouraging and essential progress in establishing and understanding the biological basis of the diseases, but how does it help us achieve the goal of moving rapid, accurate diagnostics and effective medicines forward? Well, understanding how we do this might require an appreciation of our technique first, but feel free to skip the explanation in the box below if you just want to see the results.

Traditional analysis methods like GWAS and polygenic risk scores are based on finding single SNPs or genes that have a significant association with disease across a whole population. This assumes that these single SNPs or genes explain some or all of the disease in the whole population.

The biology of complex diseases is not that simple. It is based on the effects of interactions between multiple genes and other factors, which are incredibly hard to find. Our key differentiator is that we can find such interactions – where combinations of 3, 4, 5 or more SNPs are found together to have a strong effect on disease. This enables us to find many more SNPs and genes in much smaller populations, including subgroups whose members share a specific cause of disease. Our results also reproduce much better across different populations.

You can visualize this simple but crucial difference in disease biology using simple sums. Assuming there are 4 things (we'll just use numbers) associated with a disease, traditional methods assume they are independent and their impact on disease is just a simple sum of the numbers (in any order):

Traditional (linear additive) biology *(simply adding 4 features together as independent numbers)*

```
1 + 2 + 3 + 4      = 10
```

In reality, genes and mutations interact with one another, sometimes amplifying each other's effects and sometimes reducing them. This means the way the things combine is critical, leading to more complex calculations with unpredictable results that you'd never see using the traditional approach:

Complex (non-linear) interactions in biological networks *(showing different effects of amplification and reduction between the component features of the combinations in different biological contexts)*

```
 1 – (2 x 3 x 4)  = -23
(1 + 2) x (3 x 4) = +36
```

As one of these results is positive, it might act to cause the disease, while the other negative combination might instead act to prevent the disease. Both scenarios may even involve the same gene in different contexts – yes, biology is that complex. For the poker players among you there is another explanation of this same principle of the importance of combinations here.

When we started thinking about using our results as diagnostics, we first did a very simple experiment based on the results shown in Figure 1. In the original ME/CFS analysis we found 199 SNPs, which occurred in 84 unique combinations of 3, 4, and 5 SNPs (triplets, quadruplets and quintuplets) – we call these combinations disease risk signatures.

When we simply counted how many times those 84 disease risk signatures were present in each ME/CFS patient, we got the distribution on the left of Figure 3, with patients averaging just over 4 risk signatures, but some having up to 24. In healthy people we saw far fewer, with most healthy people having 0 or just 1 risk signature. The more the shapes of these distributions

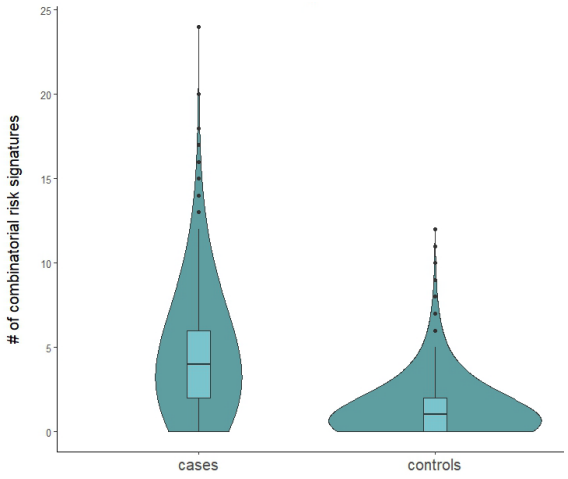

Figure 3. Distribution of disease risk signatures for ME/CFS in patients (left) vs healthy controls (right) from UK Biobank

differ, the better the prediction will be for whether a person has high genetic risk of getting ME/CFS vs a low genetic risk.

As you can see the shapes look very different for ME/CFS, which means that we potentially have a useful genetic risk test. To be clear, this is not the same as traditional diagnostic tests, which measure something affected by a disease to test whether that disease is active. Our tests provide complementary information that in some applications can be more useful than traditional diagnostics.

We can evaluate a person's disease risk using a rapid, cheap and non-invasive test. Using a buccal swab (like a COVID-19 cheek swab) and a cheap and widely available genotyping array (like UK Biobank used) we can identify which of the genetic disease risk (and protective) signatures that we've found for a disease a person has. We use this information to predict their risk of disease and also the genes that are contributing to that risk. This is critical, because knowing which genes and mechanisms are likely causing the disease in a patient means that we can start to find drugs that are most likely to be effective for them, even in some cases before the worst symptoms occur.

We call these new tests Mechanostics® and they're really important for the next steps of the story.

What Comes Next?

PrecisionLife is collaborating with Metrodora Foundation and its Institute, based in Salt Lake City, Utah. Metrodora was set up to bring researchers, doctors, and patients together to rapidly improve diagnostics, treatments, and outcomes in complex conditions (like ME/CFS and long COVID) that are underserved by current pharmaceutical and healthcare systems.

PrecisionLife is running a series of clinical studies, on 1,000 patients - 500 with ME and 500 with long COVID. The MELO trial at Metrodora has two aims. The first is to evaluate the diagnostic accuracy of the Mechanostic tests and their ability to identify people at high genetic risk of the diseases.

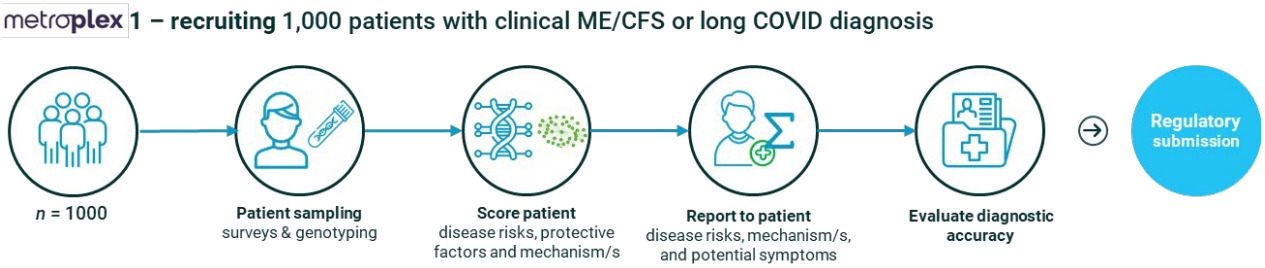

Figure 4. Arm 1 of the MELO clinical study for evaluating the ME and long COVID Mechanostics' predictive value

We have started returning results to Metrodora patients based around the specific disease risk signatures a patient has and therefore the symptoms they are most likely to experience. These reports have a lot of genetic and clinical detail but also a simple summary view of a patient's genetic risk for each of the high-level mechanisms. As shown in Figure 5, different patients present with quite different genetic risks underpinning their disease, and they and their clinicians tell us that they recognize themselves and their symptoms in the results of these tests.

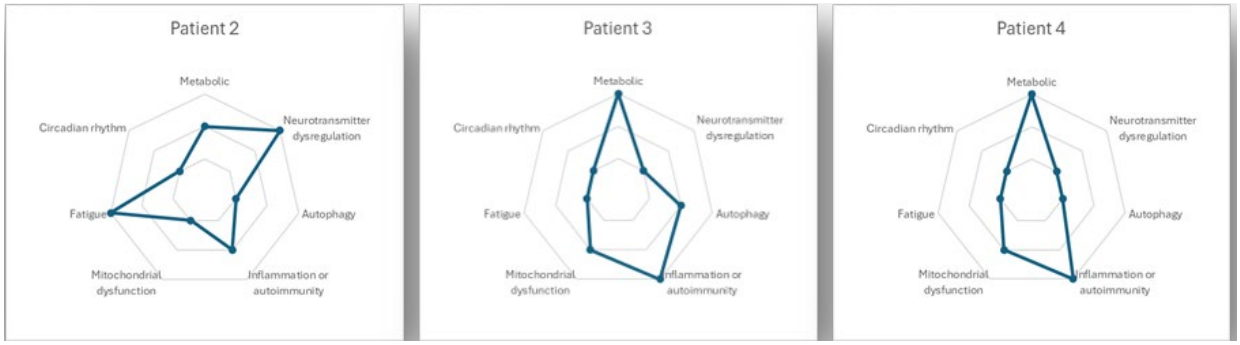

Figure 5. The clinical report's summary view showing the major disease mechanisms in ME/CFS for 3 individual patients

The second trial arm is slightly more complicated. It brings all the previous work together to find new repurposing treatments for ME and long COVID as quickly and cost-effectively as possible. We're recruiting patients from the main Mechanostics trial who have specific mechanisms driving their disease, where we have already identified that there is a safe, well-tolerated generic drug that might help patients who have that disease mechanism (i.e., a 'green drug' that can help 'green patients').

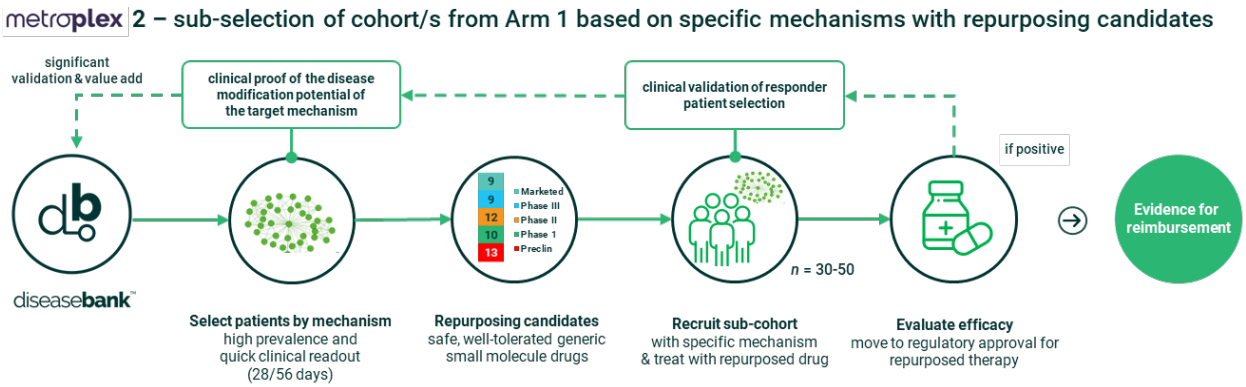

Figure 6. The design of the drug repurposing clinical efficacy trial recruiting at Metrodora

This is very similar to the design of the UK's RECOVERY trial[12,13] run in the NHS during the COVID-19 pandemic. RECOVERY quickly identified that a cheap, globally available steroid, dexamethasone, transformed the survival of many severe COVID-19 patients. This generic drug became standard of care for patients in intensive care almost overnight and has since saved over a million lives.

Our trial goes a stage further in one important respect – instead of working from clinicians' intuition about individual drugs, we have a clear genetic rationale for all of the drugs that might work in the disease and a test that can show which patients are most likely to benefit from them. PrecisionLife has analyzed all the pharmaceutical and biotech companies' pipelines to find drugs that are on market or in development against the novel disease targets we've identified, that may have the potential to benefit patients with specific mechanisms who don't currently have any effective options.

We're working closely with clinicians and patients to choose the first set of drug repurposing candidates. These are generic small molecule drugs, which have come off-patent, so they have a long history of safe prescription, are easy to dose and cheap to source. As small molecules they can be given orally as pills (rather than expensive injections), and we're starting with mechanisms that we believe have the potential to impact symptoms and show clinical benefit within 1-2 months.

We have more drug repurposing candidates than we can run at Metrodora though, and would benefit from additional support and clinical partnerships to do more of these trials in the UK, US and beyond, not just in ME/CFS and long COVID but across many other underserved diseases.

If these trials are successful, it will demonstrate to pharmaceutical companies that ME/CFS and long COVID are not too complex to make effective drugs. These are huge unserved markets, and **having the clinical validation that we can find disease causing genes, that drugs targeting these genes can bring clinical benefit to patients, and that we can even predict which patients are likely to benefit from these drugs, would be a game changer for pharma companies' appetite for investment in these diseases**. Successful validation would provide unprecedented confidence that we have the tools to accelerate and derisk novel disease biology, creating a very large and currently wide-open market.

Just One More Thing… Actively Protective Biology

We have one more exciting new opportunity that we're working on. Usually when we analyze patient datasets, we're looking for genes that are associated with (and ideally causing) a disease. This signal is the easiest to pick up – at a very basic level it simply relies on being able to find which mutations show up significantly more in patients with a disease than in people who don't have the disease.

[12] https://www.ukri.org/who-we-are/how-we-are-doing/research-outcomes-and-impact/mrc/recovery-trial-identifies-covid-19-treatments/

[13] https://www.recoverytrial.net/

Our deeper understanding of disease risk enables us to go much further. We can build a highly predictive risk score that enables us to identify the lifetime risk an individual has for getting a particular disease. Unlike the quite controversial polygenic risk scores, which are based on population averages and contain limited insights into the mechanisms driving disease at a patient level, ours are based on the deep insights into disease and the prediction of mechanisms at work in individual patients as shown in Figure 1.

We use this to find people who are completely healthy even though they have a very high-level of disease risk factors. We look for people with very high genetic risk, who have experienced all the disease triggers, and are as old as possible so they've had every chance to be diagnosed, but nonetheless they don't get the disease.

We have every reason to expect that these 'protected' people should have the disease, so we can speculate that certain factors are working behind the scenes to prevent them from showing symptoms. We can find out what makes these 'protected' people unique by comparing which genes and mechanisms show up significantly more often in them than in patients with the disease. This is an enriched reversal of the usual way we find disease risk genes.

We have a new paper just released on this, which highlights an 'actively protective' study run in ME/CFS[14]. In spite of working in very small datasets, **we found many protective signatures, mapping to 9 protein-coding genes that we believe merit further study as they may have a protective effect resisting the onset and progression of ME/CFS symptoms.** Many of these genes' functions are consistent with alleviating known ME/CFS disease risk mechanisms, e.g., in insulin signalling, stress response and autoimmunity. In other diseases the protective signals we've identified have been associated with mechanism that are subject of approved drugs.

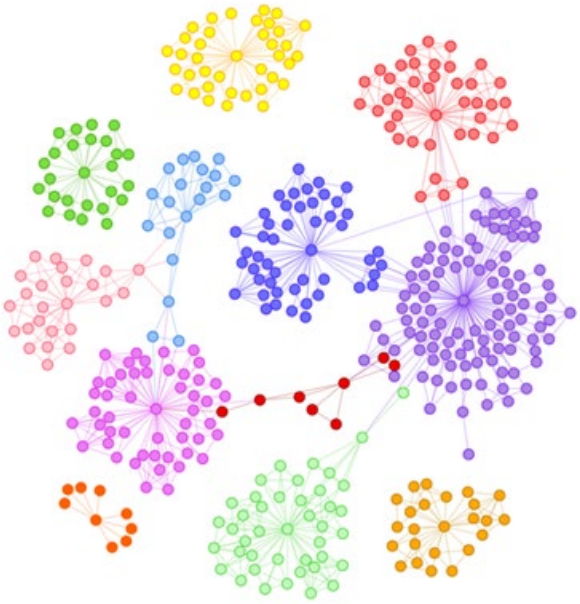

Figure 7. SNPs mapping to 9 genes identified as actively protective factors for ME/CFS by PrecisionLife's combinatorial analytics platform. Each circle is a SNP-genotype, edges connect SNP-genotypes that are co-associated in 'protected' controls, and colours represent distinct 'protective' sub-types that are not mutually exclusive in individuals (i.e., a person may benefit from more than one such protective mechanism).

These new 'actively protective' genes are a completely new class of potential drug targets. In the future these may open up opportunities for new therapies that could have a prophylactic benefit for many people who have high-risk of a specific form of a disease. They could work to reduce the disruption to normal metabolism that causes disease. This is analogous to statins, which provide protection by reducing cholesterol levels, which in turn mitigates against the major causes of cardiovascular risk.

We can also speculate that in the future some of these protective genes might turn out to be good targets for therapeutic mRNA vaccines. mRNA vaccines were used in COVID-19 to turn the body's muscle cells into factories producing parts of the viral spike proteins to raise an 'artificial' immune response. In this case we'd be using them to stimulate production of the fully-formed actively protective protein products of the selected genes.

Switching on or turning up these beneficial protective mechanisms may well benefit a much wider range of people than just those who have active disease, and these may form some important new preventative medicine tools that will be essential to delivering the healthy aging and 'increasing healthspan' agenda.

[14] J Sardell, S Das, K Taylor, C Stubberfield, A Malinowski, M Strivens, S Gardner, Actively Protective Combinatorial Analysis: a Scalable Novel Method for Detecting Variants that Contribute to Reduced Disease Prevalence in High-Risk Individuals, *Artificial Intelligence in the Life Sciences*, 2025, 100125,

The Year/s To Come…

This is very encouraging progress, but it's only the start. Use of new combinatorial analysis methods on larger genetic datasets has started to yield insights into the diseases and their mechanisms, that are transforming our understanding of the diseases and how they affect patients. We're already using these genetic insights to test better diagnostic tools and identify existing drugs that we predict might prove helpful for some patients. But genetics alone cannot take us all the way to providing real-world benefit at the scale required by patients, healthcare systems and nations. To do this we need:

- **To test more drug repurposing candidates and get more effective therapeutic options, including combination therapies based on genetic evidence showing which patients are likely to benefit, we need to scale the number and cadence of clinical trials, and bring some of this work into the NHS.**
- **To evaluate the progression of the diseases, show why patients have different symptoms and severities, and measure response to new therapies, we need to add other 'multi-omic' molecular data including proteomics, immunophenotyping and epigenetics to our patient cohorts.**
- **To more accurately understand the presentation of the diseases and the effects of potential therapies we also need much higher resolution longitudinal information on patients, which the UK is uniquely positioned to provide. We can build a UK digital cohort whose GP practice-linked electronic health record data, linked to NHS lab test results, prescription data and DNA, with self-reported and wearable data documenting their symptoms. With such a resource, we would have confidence that we can go well beyond what's already been achieved and make a huge real-world difference for the NHS and its patients, and the tens of millions of ME and long COVID sufferers around the world.**

There's still a huge amount of work to do, and it will involve all of us, from patients and their advocates, disease charities and foundations, data collections and disease studies, to innovators like PrecisionLife, healthcare systems, regulators, biopharma companies, investors, government, and right back to patients and their clinicians. We still know less than we want to, but it does feel like we have many more ways forward and new opportunities ahead of us than we did even just 12 months ago.

Partners

precisionlife improving health | for everyone

action for m·e

THE UNIVERSITY of EDINBURGH

UKRI Innovate UK

metrodora

Datasets Used

biobank uk

Sano

DecodeME

All of Us RESEARCH PROGRAM

The research described in this article has been conducted using data from UK Biobank, Sano Genetics' Long COVID GOLD study, DecodeME and the All Of Us Research Program.

These projects have been funded by PrecisionLife Ltd, Innovate UK, and Metrodora Foundation, which is supporting the US clinical trials described.

We gratefully acknowledge all of the supporters and participants in these studies for their generous and active contributions, without whom this research would not have been possible. Special thanks also to the PrecisionLife team whose dedication and innovation have made these results possible.