**////Title: Rethinking Myalgic Encephalomyelitis/Chronic Fatigue Syndrome Using Machine Learning**

**////Stand-first**:

Dr Brett Lidbury from the Australian National University worked with colleagues to utilise machine learning techniques in a new strategy to identify biomarkers that could be used to help diagnose myalgic encephalomyelitis/chronic fatigue syndrome in patients. Their work represents a significant step forward in understanding, diagnosing and treating this challenging condition, particularly in relation to pathology, the results of which form a routine but important part of general health assessment.

**///Bodytext:**

Myalgic encephalomyelitis (ME for short and also known as chronic fatigue syndrome or CFS), is a debilitating condition that remains poorly understood by clinicians and researchers alike. ME/CFS is broadly defined as a group of symptoms that include extreme tiredness that doesn’t improve with rest lasting at least 6 months, memory and concentration problems, and muscle pain.

…

Currently, there are no specific medical tests that can be used to confirm ME/CFS, and diagnosis is made based on the presence of symptoms and the exclusion of other potential conditions. Most commonly affecting adults aged in their 20s–40s, ME/CFS is associated with a particularly poor quality of life with patients often being house-bound and unable to work full time.

…

A key goal for researchers studying ME/CFS is to identify biomarkers associated with the condition that can be used to diagnose, characterise and monitor the disease. Biomarkers are molecules in the body that might be raised or depleted in certain disease states and can be measured using routine methods like blood tests. A common example is high cholesterol being a biomarker for an increased risk of heart disease.

Dr Brett Lidbury from the Australian National University and his colleagues used machine learning techniques in a novel approach aiming to identify biomarkers that could be used to predict ME/CFS. As well as ME-associated patterns among patient pathology results, they found that the biomarker activin B, previously identified as a candidate in a 2017 pilot study by Dr Lidbury and his colleagues, could be used as a diagnostic marker and to predict symptom severity.

…

In the current study, Dr Lidbury and his colleagues recruited 80 participants with ME/CFS and 17 healthy control participants via the CFS Discovery centre in Victoria, Australia. Participants in the ME/CFS groups all fulfilled the International Consensus Criteria, which is a checklist used by clinicians to guide the diagnosis of ME/CFS.

All participants completed a standing test to determine the severity of their ME/CFS, and blood and urine samples were collected for biomarker investigations. These included serum activin B, urinary creatinine, serum urea, mean corpuscular haemoglobin, alkaline phosphatase, and the total lymphocyte count, among others typically used by GPs when establishing blood test profiles.

Patients with ME/CFS generally have difficulty standing for long periods of time, so the standing difficulty test is used to calculate a ‘weighted standing time’ score. A difficulty score of 0 means the patient has no difficulty standing upright for 20 minutes, whereas a score of 10 or greater reflects difficulty to maintain an upright posture. A score of 14 was given to patients who couldn’t stand for longer than 10 minutes and 12 was given to those who could stand for at least 10 minutes, but not the entire 20 minute test period. The weighted standing time score was then used to divide the cohort of ME/CFS patients into three categories, which were mild, moderate and severe. These severity categories were used during statistical analysis when looking for trends amongst the biomarkers recorded from the blood and urine samples taken.

Machine learning is a form of artificial intelligence that focuses on using algorithms that can classify input or identify trends through exposure to more data. As machine learning algorithms tend to ‘learn’ by being trained on data, they are well placed to handle large volumes of data that have traditionally been difficult for researchers to handle manually. They are also useful for identifying trends in complex datasets that incorporate many different variables and allow data-based predictions of an outcome.

Over the last 5 years, machine learning has been increasingly applied to medical studies to demonstrate enhanced handling of large patient datasets, both old and new, to provide new insights into conditions. Random forest, as used here by Dr Lidbury and his colleagues, is a machine learning technique used to predict outcomes from datasets using multiple decision trees.

In this work, Dr Lidbury and his colleagues provided two critical results. The first was that they were able to use levels of serum activin B from blood samples to successfully identify ME/CFS participants from healthy controls.

The second was that from their random forest analysis, urinary creatinine, alkaline phosphatase and activin B could all be used to predict the severity of ME/CFS in participants as defined by the statistically derived standing scores. When using the biomarkers in conjunction with each other –specifically the addition of activin B – the prediction error for different ME/CFS severity classes in the random forest analysis could be reduced. Overall, the study confirmed the utility for activin B to impact the separation of ME patients experiencing mild to moderate symptoms, as reflected by weighted standing time. This is useful because although severe disease is normally easy to define based on clinical symptoms alone, milder forms of disease are not.

These findings are important. Not only do Dr Lidbury and his colleagues provide a practical method for achieving early and readily accessible diagnosis [i.e., screening tools for primary health professionals via pathology (blood test) results] and severity prediction of ME/CFS, they also provide some insight into the biological mechanisms of the disease process. In conjunction with studies in adjacent fields like immunology, biology and metabolomics, the trends in biomarkers observed here could be used to identify factors like impaired metabolomic pathways or the impact of the condition on immune response in future work. Further, the study also demonstrates the applicability of machine learning techniques to clinical data and provides a precedent for optimised algorithm development in the future.

Dr Lidbury and his colleagues identified some considerations in the current study that should be addressed in future work. The first is the trend they observed in activin B levels. In their previous 2017 study, they found that activin B levels were higher in their ME/CFS group compared to the control group. However, in the current work, they found the opposite, and this requires exploration.

The second was the wide spread of disease severity in their cohort, which led to overall smaller sample sizes in each category. This prompted their decision to group mild ME/CFS with the healthy control group during statistical analysis to increase the size of the control cohort. Again, conducting larger multi-centre studies with a larger number of participants will be useful in clarifying the researchers’ findings to date.

Importantly, Dr Lidbury and his colleagues have already made significant steps forward in determining the potential of activin B as a biomarker for ME/CFS, paving the way forward for more extensive work in the field.

This SciPod is a summary of the paper ‘Rethinking ME/CFS Diagnostic Reference Intervals via Machine Learning, and the Utility of Activin B for Defining Symptom Severity’, from the journal Diagnostics. DOI: https://doi.org/10.3390/diagnostics9030079

For further information, you can connect with Brett Lidbury at brett.lidbury@anu.edu.au