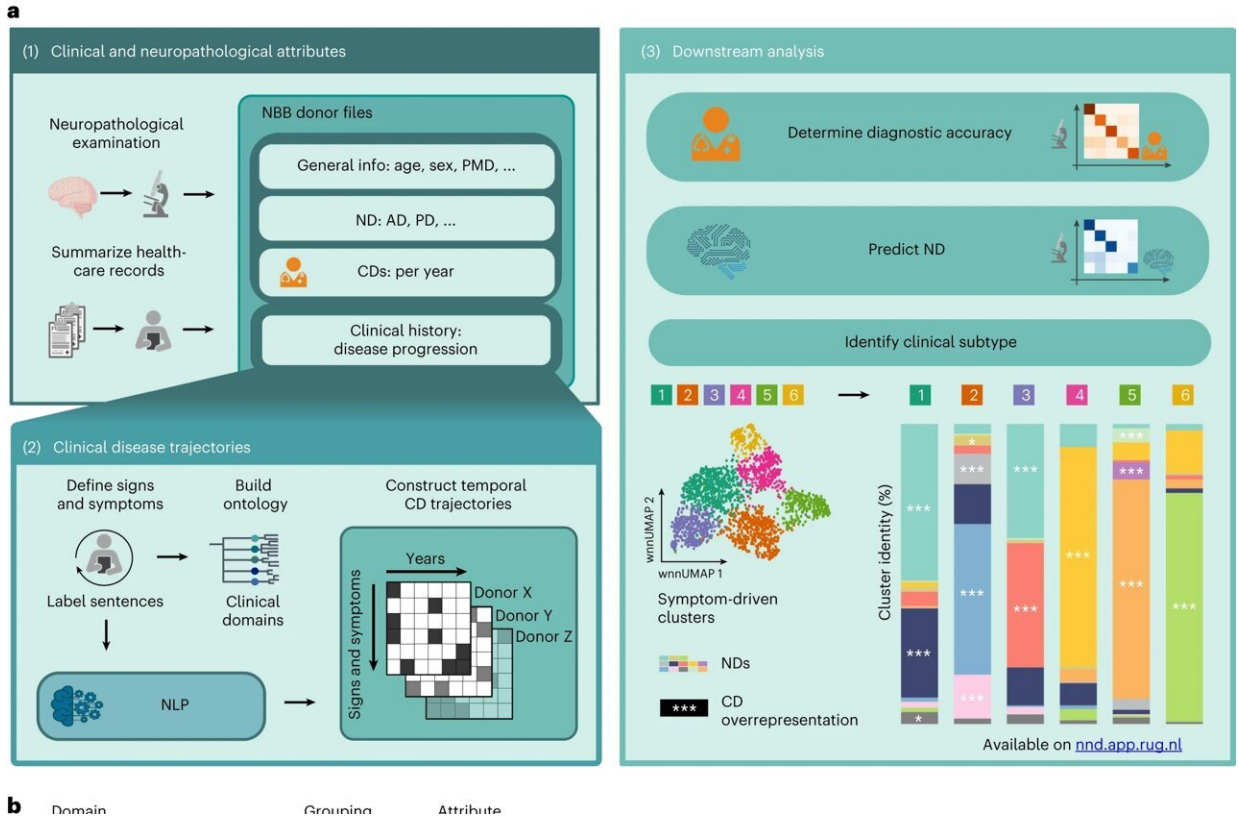
AI language model provides new insights into the development of brain diseases

March 12 2024



Introduction to the project. a, Workflow of the project describing the different data types in the NBB donor files (i), the processing of the clinical history data resulting in clinical disease trajectories (ii) and downstream analyses (iii). b, Clinical attributes (signs and symptoms), their domains, and groupings, including colors and illustrative brain icons. Credit: *Nature Medicine* (2024). DOI: 10.1038/s41591-024-02843-9

A new AI language model identifies clinical symptoms in medical summaries and links them to brain tissue from donors of the Netherlands Brain Bank. This yields new insights into the development of individual disease progression and contributes to a better understanding of common misdiagnoses of brain diseases. The model may, in the future, assist in making more accurate diagnoses.

In many brain diseases, the underlying molecular mechanisms are often poorly understood, making it challenging to develop new treatment options. Investigating these molecular mechanisms is additionally challenging because the relationship between actual tissue abnormalities and the patient's symptoms is often highly complex.

Some symptoms, for example, occur in multiple conditions, and the clinical picture can vary significantly from patient to patient, resulting in a substantial percentage of misdiagnoses (up to 30 percent). Insights gained from a newly developed AI language model may potentially change this scenario in the future.

At the Netherlands Brain Bank, brain tissue from 3,042 brain donors with a wide range of different brain diseases is stored. What makes the Netherlands Brain Bank unique is that, in addition to the tissue, they have also documented the medical history and the symptoms reported by the donors. However, this wealth of data was not quantifiable because it was transcribed in a text format, making it difficult to analyze and process.

Language model

Inge Huitinga and her team at the Netherlands Institute for Neuroscience joined forces with Inge R. Holtman and her team at the University Medical Center Groningen to unlock this information using a new AI language model. This classification model enables the analysis of the text

in medical records and the detection of predefined symptoms. Additionally, they developed a second AI prediction model to make actual diagnoses based on the clinical picture.

Holtman says, "First, the records had to be thoroughly examined to identify symptoms that regularly occur in donors with different brain diseases. We eventually identified 90 different symptoms in five different domains: psychiatric symptoms (such as depression and psychosis), cognitive symptoms (such as dementia and memory problems), motor issues (such as tremors), and sensory symptoms (such as feeling things that are not there)."

"We then manually labeled 20,000 sentences to train the classification model."

The final model ultimately determined which symptoms occurred annually for all donors. It was observed that the prediction model was quite effective in making accurate diagnoses but fell short in rare disorders. When analyzing the diagnoses made by the prediction model, a subset of donors emerged who had been incorrectly diagnosed. It turned out that the doctor had also misdiagnosed a considerable number of these donors during their lifetime.

Subtypes

Holtman says, "It seems that there is a group of people suffering from a certain condition, such as Alzheimer's disease, but exhibiting symptoms more reminiscent of Parkinson's disease. Or a subtype of Frontotemporal Dementia manifesting as Alzheimer's disease."

"It is often challenging to diagnose these groups properly, which makes sense since these individuals show a clinical pattern that does not align with their condition. We strive to continuously improve the prediction

model, hoping to make diagnoses of brain diseases more accurate."

Huitinga explains, "Understanding individual factors contributing to symptoms in brain diseases is crucial, as the reality is that many people have a combination of different conditions. Molecular markers to guide treatment are the future. Our ultimate goal is to develop a molecular atlas of symptoms of brain diseases. Such an atlas precisely shows which cells and molecules in the brain change with symptoms such as anxiety, forgetfulness, and depression."

"We expect the impact of this molecular atlas to be enormous. When we map out the molecular changes, we hope to identify the first biomarkers that can predict the correct diagnosis during a person's lifetime. This opens doors to the development of new therapies. We are laying the foundation."

The study is published in the journal *Nature Medicine*.

More information: Nienke J. Mekkes et al, Identification of clinical disease trajectories in neurodegenerative disorders with natural language processing, *Nature Medicine* (2024). DOI: 10.1038/s41591-024-02843-9

Provided by Netherlands Institute for Neuroscience