Detection of cognitive deficits years prior to clinical diagnosis

across neurological conditions

Xin You Tai, MD, DPhil; Sofia Toniolo, MD, DPhil; David Llewellyn, PhD; Cornelia M. van Duin, PhD; Masud Husain, MD, DPhil; Sanjay Manchikanti, MD, DPhil

1 Nuffield Department of Clinical Neuroscience, University of Oxford, Oxford UK
2 Division of Clinical Neurology, John Radcliffe Hospital, Oxford University Hospitals Trust, Oxford, UK
3 College of Medicine and Health, University of Exeter, Exeter, UK
4 Alan Turing Institute, London, UK
5 Nuffield Department of Population Health, University of Oxford, Oxford, UK
6 Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Oxford, UK
7 Department of Experimental Psychology, University of Oxford, Oxford, UK

*Denotes equal contribution

Corresponding author: Xin You Tai
Corresponding author’s address: Level 6, West Wing
John Radcliffe Hospital
University of Oxford
Corresponding author’s e-mail: xin.tai@ndcn.ox.ac.uk

Number of figure and tables: 4
Word count: 3339 (main text) and 360 (abstract)

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Importance: Understanding the cognitive trajectory of a neurological disease can provide important insight on underlying mechanisms and disease progression. Cognitive impairment is now well established as beginning many years before the diagnosis of Alzheimer’s disease, but pre-diagnostic profiles are unclear for other neurological conditions that may be associated with cognitive impairment.

Objective: Compare pre-diagnostic and post-diagnostic cognition and global brain volume in ischaemic stroke, focal epilepsy, Parkinson’s disease, multiple sclerosis, motor neurone disease (amyotrophic lateral sclerosis) and migraine using time-to-diagnosis and time-from-diagnosis data in relation to time of assessment.

Design: Analysis the prospective UK Biobank cohort with study baseline assessment performed between 2006-2010 and participants followed until 2021.

Setting: Multicenter, population-based study.

Participants: Sample of 497,252 participants, aged between 38 and 72 years, at baseline with an imaging sub-sample of 42,468 participants.

Exposure: Participants with each neurological condition were compared to a healthy control group.

Main outcomes and measures: A continuous measure of executive function and magnetic resonance imaging brain measures of total grey matter and hippocampal volume.

Results: Of the 497,252 participants (226,206 [45.5%] men, mean [SD] age, 57.5[8.1] years), 12,755 had ischaemic stroke, 6,758 had a diagnosis of focal epilepsy, 3,315 had Parkinson’s disease, 2,315 had multiple sclerosis, 559 had motor neurone disease and 18,254 had migraine either at study baseline or diagnosed during the follow-up period. Apart from motor neurone disease, all conditions had lower pre-diagnosis executive function compared to controls (assessment performed median 7.4 years before diagnosis). Participants with focal epilepsy and multiple sclerosis showed a gradual
worsening in executive function up to 15 years prior to diagnosis, while ischaemic stroke was
characterised by a modest decline for a few years followed by a substantial reduction at the time of
diagnosis. By contrast, participants with migraine showed improved post-diagnosis cognitive scores.
Pre-diagnosis MRI grey matter volume was lower than controls for stroke, Parkinson’s disease and
multiple sclerosis (scans performed median 1.7 years before diagnosis), while other conditions had
lower volumes post-diagnosis.

**Conclusion:** These cognitive trajectory models reveal disease-specific temporal patterns, including a
long cognitive prodrome associated with focal epilepsy and multiple sclerosis. The findings may help
to prioritise risk management of individual diseases and inform clinical decision-making.
Key Points

Question

What is the pre-diagnosis cognitive profile across neurological conditions of ischaemic stroke, focal epilepsy, Parkinson’s disease, multiple sclerosis, motor neurone disease and migraine?

Findings

This cohort study of 495,149 participants, aged between 38 and 72 years, identified gradual worsening of cognition up to 15 years prior to clinical diagnosis in participants with focal epilepsy and multiple sclerosis while stroke was associated with a modest decline for a few years followed by a substantial reduction at the time of diagnosis. Migraine was associated with post-diagnosis improvement in cognitive scores.

Meaning

Cognitive trajectory models identify disease-specific temporal patterns that may help to prioritise risk management of individual diseases and inform clinical decision-making.
Introduction

The cognitive profile of a neurological disease across time can provide important insights into underlying mechanisms, response to treatment, disease progression and prognosis. In Alzheimer’s disease, cognitive impairment may occur decades prior to the diagnosis which reflects ongoing accumulation of neurodegenerative protein and associated cellular dysfunction and death. Following diagnosis, comparison of cognitive trajectories provides an index of response to treatment such as cholinesterase inhibitors and the recently described amyloid immunotherapies. In other neurological conditions that may cause cognitive impairment, however, pre- and post-diagnosis cognitive profiles are largely unknown.

Longitudinal evaluation is crucial to map cognitive trajectories. Once a diagnosis has been given, individuals with a condition can be followed up with repeat cognitive measurements to understand the post-diagnosis cognitive changes. Delineating cognition before the disease is diagnosed, however, is less straightforward. The current pre-diagnosis cognitive model of Alzheimer’s disease has been classically derived from autosomal dominant Alzheimer’s disease, leveraging families with offspring that have 50% chance of developing the disease. This cognitive model is corroborated by longitudinal studies of individuals with mild cognitive impairment who progress to develop Alzheimer’s disease.

For other neurological conditions, pre-diagnostic assessment is often not feasible, and even if it is, longitudinal evaluation is resource-intensive and logistically difficult. Therefore, many important questions remain unanswered. For example, what is the potential and timeframe of post-stroke cognitive recovery? Many studies describe post-stroke cognitive impairment but few report data from multiple time points or performance after one year. Do individuals with focal epilepsy have cognitive impairment prior to starting anti-epileptic medications? One group identified worse memory scores present in new cases of focal epilepsy before treatment initiation, but precisely when these difficulties begin is unclear. Similarly, what is the cognitive impact of multiple sclerosis
and does this improve with treatment? Recent evidence suggests that cortical lesions exist in the
early pathogenesis of some people with multiple sclerosis which may affect cognitive function.\textsuperscript{16}
However, there is little corroboration outside of small studies. One important related consideration
is that cognitive decline may be coupled with reduction in brain volume observed in these
conditions.\textsuperscript{17–19}
In this study, we use a data-driven approach to understand both pre- and post-diagnosis cognitive
profiles and neuroimaging data from the UK Biobank prospective cohort across a range of
neurological diseases. Specifically, in relation to date of diagnosis, we model the time-course and
magnitude of cognitive impairment across individuals with ischaemic stroke, focal epilepsy,
Parkinson’s disease, multiple sclerosis, motor neuron disease (amyotrophic lateral sclerosis) and
migraine, compared to healthy controls. We hypothesized that conditions such as focal epilepsy and
Parkinson’s disease would be associated with worse cognition many years before the diagnosis is
made, whereas individuals with stroke may experience cognitive impairment at the time of
diagnosis, as this reflects an index event of brain injury. We also compare the magnitude of cognitive
deficit across these neurological conditions to provide a better understanding which may inform
clinical decision-making on risk reduction strategies and long-term management.

Methods
This study examines the UK Biobank, a population-based cohort of over 500,000 participants aged
38-72 years who underwent physiological measurements, cognitive testing and provided biological
samples at one of 22 centres across the UK between 2006 and 2010.\textsuperscript{20} A subset re-attended for brain
imaging between 2014 and 2020.\textsuperscript{21} All participants provided written informed consent. UK Biobank
received approved from the North-West Multi-centre Research Ethics Committee (REC number
21/NW/0157).
The primary study objective was to investigate cognition across individuals with ischaemic stroke, focal epilepsy, Parkinson’s disease, multiple sclerosis, motor neurone disease and migraine. These diagnoses were chosen as common, acquired neurological conditions that may present to a general neurology clinic. Established cases at study baseline assessment (post-diagnosis cases) as well as new diagnoses during longitudinal follow-up were identified from hospital inpatient records, coded using the International Classification of Diseases (ICD) ICD-9 and ICD-10 codes, or from death register linkage data as an underlying or contributory cause (UK Biobank codes are found in Table S1).

Secondary analysis considered all-cause dementia diagnosed after baseline assessment, which provided a benchmark with an expected cognitive decline following diagnosis, and sub-types including Alzheimer’s disease, vascular dementia and frontotemporal dementia. There were few cases of dementia at baseline assessment (N = 120), which were excluded due to the direct association with cognitive impairment. Additional exclusion criteria included other neurological conditions such as history of CNS infection, encephalitis, meningitis, haemorrhagic stroke, genetic epilepsies, previous subdural or subarachnoid haemorrhage.

**Cognitive testing**

We analysed data from five computer-based cognitive tasks of working memory or speed of processing as in a previous reports focusing on executive function in the UK Biobank population, and used the first available timepoint data. The tests were a pairs-matching and snap reaction time, trail-making, tower-rearranging, and symbol-digit substitution tasks, described in detail elsewhere with reliability and retest-effects previously assessed.

**Pre-diagnosis and post-diagnosis cognitive profiles**

Cognitive profiles were created in relation to the date of diagnosis whereby post-diagnosis cognition represented all the participants with a condition who were diagnosed prior to baseline study.
assessment for that individual. For example, the cognitive performance of a participant who was diagnosed with multiple sclerosis 20 years prior to entering the UK Biobank study represented a 20-year post-diagnosis cognitive score for that condition. By contrast, two-year pre-diagnosis cognition reflected cognitive scores at baseline assessment for participants who would (subsequently) be diagnosed with the condition two years later during the study follow-up period.

Main covariates

All full models were adjusted for age (continuous), sex (female vs male), education (categorised as higher [college or university degree or other professional qualification], upper secondary [second or final stage of secondary education], lower secondary [first stage of secondary education], vocational [work-related qualifications], or other), socioeconomic status (categories derived from Townsend deprivation index quintiles 1, 2 to 4, and 5).

Brain imaging variables

Magnetic resonance imaging (MRI) data were acquired on a Skyra 3T scanner (Siemens; Munich, Germany) including high-resolution, T1-weighted, three-dimensional magnetisation-prepared gradient echo structural images and T2-weighted fluid-attenuated inversion recovery images. Full imaging protocols and processing pipeline have been previously described. We examined a global brain measure of total grey matter volume, useful to assess wide-spread change, and the total hippocampal volume which is related to cognition and relevant for several of the diseases such as epilepsy and stroke. Median absolute deviation was used to exclude outliers, and volumes were adjusted for potentially confounding baseline measures of age, age squared, head size, and imaging site.

Statistical analysis
Confirmatory factor analysis (CFA) was performed on cognitive variables to produce a continuous, summary latent measure of working memory and reaction time which we termed “Executive Function” for simplicity (this method has been previously described in detail).\textsuperscript{23, 31} Estimating a latent variable has the methodological advantage of controlling measurement error that can artificially reduce the relationship between measured variables in standard univariate analyses.\textsuperscript{32} Missing cognitive data were estimated using full information maximum likelihood, which gives unbiased parameter estimates and standard errors.

The relationship between Executive Function across age was first examined for each group - ischaemic stroke, focal epilepsy, Parkinson’s disease, multiple sclerosis, motor neurone disease and migraine. We then controlled for age in two different ways for robustness: using a residual-based, sliding window approach with fixed age-quantile widths moved along the age distribution (using a 20% gaussian kernel and described previously,\textsuperscript{23, 33, 34} code: conditionalPlot.m available: \url{https://osf.io/vmabg/}). This has the methodological advantage of modelling non-linearities in the data and is not constrained by a pre-specified model. The second method included age as a covariate in a general linear model along with other baseline characteristics. The difference in Executive Function between pre-diagnosis cases, post-diagnosis cases and the control group was compared using ANOVA with post-hoc Tukey analysis to account for pairwise or multiple comparisons, p-values were two-sided with statistical significance set at p<0.05 for all analyses. Brain measures were analysed using the same methods.

The time course of cognitive scores and brain measures were visualised using the same sliding window, model-free approach with fixed quantile widths of time in relation to diagnosis. We compared this data with age-residualised cognitive scores from the control group. Motor neurone disease was not included in this visualisation due to low sample size but was analysed when comparing pre- and post-diagnosis Executive Function and brain imaging volumes. For the main exposures and covariates, there were less than 3% missing or not known data and complete case analysis was applied. New diagnoses were included if recorded between baseline until the date of
first diagnosis, death, loss to follow-up, or the last surveyed hospital admission date (March 31, 2021, for England and Scotland and Feb 28, 2018, for Wales), whichever came first. These censoring dates were recommended by UK Biobank as the data was estimated to be over 90% complete in England, Scotland, and Wales. Analyses were done in Matlab R2018a or in R version 4.0.3 using the Lavaan or survival package.

Results

The UK Biobank cohort comprised 502,536 participants at baseline. After excluding those who did not meet the inclusion criteria (N = 7216), our study included 497,252 individuals (flowchart in Figure S1). Overall, participants had a mean (SD) age of 57.5 (8.1) years and 54.5% were female. At study baseline, there were 6,352 with ischaemic stroke, 4,247 participants with epilepsy, 873 with Parkinson’s disease, 1,861 with multiple sclerosis, 70 participants diagnosed with motor neurone disease and 14,588 with migraine. The median time-since-diagnosis to baseline study assessment in these post-diagnosis cases ranged from 27.6 years in individuals with migraine to 3.3 years in individuals with MND.

Over 5,826,307 total follow-up years (median 12.0, interquartile range 11.2-12.7 years), there were 6,403 participants with a new diagnosis of ischaemic stroke, 2,511 participants with focal epilepsy, 2,442 with Parkinson’s disease, 454 with multiple sclerosis, 489 with motor neurone disease and 3,666 with migraine. The median time-to-diagnosis was 7.4 years from baseline study assessment in these ‘pre-diagnosis’ cases across all conditions. Individual condition demographic information can be found in Table 1 while histograms of the diagnosis dates in relation to baseline study assessment are shown in Figure S2.

Pre-diagnosis and post-diagnosis Executive Function across neurological conditions
A continuous cognitive function latent variable of Executive Function was estimated from five cognitive tasks of working memory or speed of processing (model and fit indices shown in Figure S3).

Age is often the most important confounder when considering cognition and, consistent with previous findings, Executive Function declined uniformly with age among all groups (Figure S4). We therefore controlled for age using a residual-based method. Average Executive Function scores were lower in individuals with an established diagnosis (post-diagnosis) compared to individuals before they were diagnosed with a neurological condition (pre-diagnosis), apart from in those with migraine and motor neurone disease (Figure 1, Figure S5 shows individual condition plots and Table S2 with post-hoc Tukey analysis results, controlling for multiple comparison testing). Further, pre-diagnosis Executive Function scores were significantly lower than controls for all conditions apart from motor neurone disease. For robustness, we used a separate general linear model to corroborate these findings and control for other covariates including sex, level of education and socioeconomic status. Similar patterns of differences were observed when comparing pre-diagnosis and post-diagnosis Executive Function scores with controls (Table S3).
Figure 1. Pre-diagnosis and post-diagnosis Executive Function across different neurological conditions

Prior to being diagnosed, participants with epilepsy, stroke, migraine, Parkinson’s disease and multiple sclerosis have a lower Executive Function (z-scored values) compared to controls, with a median time-to-diagnosis of 7.4 years. Dotted baseline represents level of Executive Function in controls. Across all neurological conditions, apart from migraine and motor neurone disease, Executive function was lower in individuals with an established diagnosis (post-diagnosis cognition) compared to participants who would be diagnosed during the follow-up period of the study. Participants with migraine had a higher post-diagnosis Executive Function compared to pre-diagnosis participants. Error bars denote standard error. Mean executive function for each condition was compared with control group, asterisk (*) below each bar denotes significant difference p<0.05, while within-condition pre-diagnosis vs. post-diagnosis executive function was compared and represented by an asterisk across condition bars using a post-hoc Tukey analysis to account for multiple comparison. PD – Parkinson’s disease, MS – multiple sclerosis, MND – motor neurone disease

Modelling cognitive trajectories across time in relation to diagnosis

Using a non-linear approach, Figure 2 shows the age-residual Executive Function scores leading up to and after diagnosis for each neurological condition (Figure 2). The upper limit of this time course reflects participants who were diagnosed many years before entering the UK Biobank study and this was different for each neurological condition. The lower limit of these cognitive trajectories
represents individuals who were diagnosed during the follow-up period of the UK Biobank study (up
to 15 years follow-up after baseline assessment).

This approach identified disease-specific patterns in cognitive profiles at the group level. Executive
Function gradually declined in participants with focal epilepsy over 10 years before the diagnosis
was made whereas participants with stroke showed Executive Function scores similar to controls up
until a few years before their event, then a modest worsening followed by a substantial decline at
diagnosis. Participants with multiple sclerosis had comparatively worse Executive Function
compared to all other neurological conditions assessed, with a gradual decline pre-diagnosis which
continued post-diagnosis. By contrast, participants with migraine had lower Executive Function
before their diagnosis which improved to baseline control level after the diagnosis was made. An
initial plateau in Executive Function was observed in individuals with epilepsy and stroke following
the diagnosis with a small improvement between 2-10 years observed in stroke but not to the level
of controls. Individuals with Parkinson’s disease showed a constant lower-than-control level of
Executive Function 10 years before their diagnosis. For motor neurone disease, there was mainly
data available leading up to diagnosis with few established cases at study baseline. There was worse
post-diagnosis Executive Function in individuals with motor neurone disease although this was
associated with large standard error due to smaller sample size.
Figure 2. Cognitive profile across time in relation to the diagnosis of disease across different neurological conditions.

Executive Function (age-residuals, z-scored values) in relation to time of diagnosis across participants with different neurological conditions represents a pre- and post-diagnosis cognitive trajectory. Black dotted line indicates the level of Executive Function in controls. Red dotted line indicates the study assessment time. Negative time values reflect individuals who will develop the disease (pre-diagnosis cases) during the study follow-up period. Participants with focal epilepsy, Parkinson's disease and multiple sclerosis show declining cognitive scores prior to diagnosis at a group level. Cognitive scores plateau after diagnosis in several conditions, including focal epilepsy, and further declines in multiple sclerosis. Executive Function in participants in migraine improves following diagnosis to a level similar to controls.
Pre-diagnosis cognitive profile in dementia and sub-types

There were 5963 participants who developed all-cause dementia during the follow-up period of the study. Executive Function declined progressively in individuals who were closer to receiving the diagnosis (Figure S6). This pattern was similar when considering those who were diagnosed with Alzheimer’s disease (2420 participants), vascular dementia (1019 participants) and other dementias (2325 participants) but was not observed in participants with frontotemporal dementia (199 participants).

Total grey matter pre-diagnosis and post-diagnosis across neurological conditions

A subset of the UK Biobank cohort underwent MR brain imaging (N=42,468) following baseline study assessment (range 3.8 to 13.8 years after, median 9.2 years). Since imaging was obtained later, the median time-to-diagnosis from imaging acquisition was lower at 1.7 years than cognitive assessment (Table 2 details the breakdown for each neurological condition). We used the same approach in modelling the pre- and post-diagnosis time-course for total grey matter, adjusted for age, across each neurological condition (time-course limited to 20 years post-diagnosis for visualisation in Figure 3). Total grey matter volume was lower than controls in individuals who have had an ischaemic stroke and those with multiple sclerosis and Parkinson’s disease, both pre- and post-diagnosis (general linear model showing statistical differences in Table S4). Participants with epilepsy showed decreasing total grey matter volume post-diagnosis but have volumes similar to controls leading up to the diagnosis while total grey matter volume in participants in migraine are comparable to controls before and after the diagnosis. The results for participants with motor neurone disease were not visualised because only 19 participants had imaging data.

The same pre- and post-diagnosis imaging analysis was performed for total hippocampal volume which showed lower pre- and post-diagnosis volume in participants with multiple sclerosis only, while ischaemic stroke and focal epilepsy was associated with lower post-diagnosis hippocampal


volume only (Table S5), suggesting that different disorders may be weighted towards affecting different brain areas.

**Figure 3.** Total grey matter volume across time in relation to diagnosis across different neurological conditions

Total grey matter (GM) volume (age-residuals, z-scored) in relation to time of diagnosis across different neurological conditions. Black dotted line indicates the total grey matter volume in controls. Red dotted line indicates time of the brain scan event. Negative time values reflect individuals who will develop the disease (pre-diagnosis cases) after the brain scan has been performed. At a group level, total grey matter volume in participants with multiple sclerosis, Parkinson’s disease and a history of ischaemic stroke are lower than controls before and after the diagnosis has been made. Participants with migraine have a similar age-corrected total grey matter volume throughout the time course. In participants with epilepsy, post-diagnosis total grey matter volume decreases more steeply than expected for age in participants who have had the disease longer.
Discussion

This study examined pre- and post-diagnosis cognitive profiles in large groups of individuals with ischaemic stroke, focal epilepsy, Parkinson’s disease, multiple sclerosis and motor neurone disease. By leveraging cross-sectional cognitive data at baseline study assessment along with time-to-diagnosis (in the future) and time-since-diagnosis (from the past), we analysed group-level cognitive ‘trajectories’ with profiles up to 15 years pre-diagnosis. Such trajectories have not been clearly described previously for several of these conditions. Our residual-based, sliding-window analysis allows non-linearities to emerge from the data which more closely approximates real-world disease characteristics rather than constraining analysis to linear or quadratic functions. All conditions, apart from motor neurone disease, showed a measurable cognitive deficit several years before diagnosis compared to controls, with multiple sclerosis having the lowest cognitive score across all conditions.

Our findings offer several potential disease-specific insights. Focal epilepsy has been closely linked to cognitive impairment\(^{36,37}\), however, this has been difficult to dissociate from cognitive side-effects of anti-seizure medication.\(^{35}\) In the present study, impaired cognition was evident over a decade prior to the diagnosis. This may reflect ongoing sub-clinical epileptic activity\(^{38}\) or associated neuropathology\(^{39}\) before two overt seizures occur which are required for diagnosis.\(^{40}\) Cognitive scores plateaued within two years following focal epilepsy diagnosis which may be due to treatment stability before further decline after 20 years. Pre-diagnosis cognitive deficits observed with multiple sclerosis may reflect emerging research on early B-cell related, cortical pathology that can occur before the characteristic white matter demyelinating lesions.\(^{16,41}\)

A sharp decline in cognition was observed around the diagnosis of ischaemic stroke, likely reflecting the index stroke event, while a modest reduction in cognitive scores in the few years prior may indicate underlying small vessel disease related to vascular risk factors.\(^{31,42,43}\) Interestingly, a small increase in cognitive scores between up to eight years following stroke diagnosis was observed which warrants further analysis and may relate to factors of stroke rehabilitation and brain recovery.
processes such as neuroplasticity. By contrast, individuals with migraine had lower cognitive scores leading up to the diagnosis that improved post-diagnosis back to a level similar to controls. This may represent generalised cognitive dysfunction of attention and working memory seen with migraine, commonly described as a 'brain fog', but with a potentially beneficial treatment effect following diagnosis. Importantly, the lack of ongoing cognitive decline is in line with the absence of underlying progressive or degenerative pathology in migraine.

The neuroimaging findings suggest that reduced global brain volume can be detected in participants with ischaemic stroke, Parkinson’s disease and multiple sclerosis several years before the diagnosis. While this is consistent with the neurodegenerative and neuroinflammatory processes in Parkinson’s disease and multiple sclerosis, respectively, this has not been clearly reported before and will benefit from a closer examination of different brain regions involved in each condition.

Several further observations can be made for each individual condition, but it is important to consider our findings in the context of the study limitations. Firstly, this is a broad overview into each condition as disease-specific characteristics have not been accounted for, such the location of demyelinating lesions in participants with multiple sclerosis. Medications may be more important for some of the studied diseases, for example, we have previously shown that the number of anti-seizure medications correlates to worse cognition likely as a marker of seizure severity rather than medication side effect. In-depth analysis in each disease is the subject of ongoing investigation. However, in our view, this general approach provides a useful and novel comparison across these conditions.

Crucially, due to the observational nature of this cohort, any associations or findings cannot be taken as causal and potential mechanistic explanations are hypothetical. Diagnosis information was based on hospital records or death certificate data and therefore delayed or incorrect diagnoses may affect the results. Nonetheless, early changes were observed up to 20 years before diagnosis for several of these conditions. Electronic records may be more accurate for certain conditions, such as ischaemic...
stroke, when there is usually a clear event that brings an individual to hospital compared to neurodegenerative conditions with a longer diagnostic journey. However, the sample size of each sub-group is an order of magnitude larger than most separate studies of each condition and will help identify small effects within the data. The cross-sectional nature of the data means that these cognitive and brain models likely capture a combination of disease trajectory over time and cohort-specific characteristics. This includes a selection bias that reflects the relatively healthy nature of the UK Biobank and therefore the effects described may underestimate the magnitude of cognitive changes should a more representative sample of each condition were considered.

Conclusion

Analysis of cognitive and brain trajectories can offer novel disease-specific insights and improve understanding of underlying pathological mechanisms. Cognitive deficits were identified over a decade before diagnosis in conditions such as epilepsy and multiple sclerosis while cognition post-diagnosis improved to a healthy control baseline in individuals with migraine but declined across other conditions. This non-linear approach using cross-sectional cognitive data with time-to-diagnosis and time-since-diagnosis is made possible with the large cohort of the UK Biobank and is a cost-effective and useful method to understand disease time-courses compared to longitudinal studies. These findings may help prioritise risk management of individual diseases and may inform screening, clinical diagnosis and long-term decision-making in the future.
Data availability

The data analysed during the current study are available from the UK Biobank
https://www.ukbiobank.ac.uk/researchers/. The variables used are detailed in Supplementary Table 1. This research was conducted using the UK Biobank resource under application 9462.

Acknowledgments: This work was funded by the Wellcome Trust (Wellcome Trust PhD clinical fellowship to XYT and Wellcome Trust Principal Research Fellowship to MH [206330/Z/17/Z]) and MRC (Clinician Scientist Fellowship to S.M. [MR/P00878/X]) and by the Oxford NIHR Biomedical Research Centre.

Author contributions: Design, analysis, manuscript writing, critical revisions (XYT); design, statistical analysis and critical revisions of the manuscript (ST, DL, CvD, MH, SM)

Competing Interests

The authors declare no competing interests.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Controls</th>
<th>Epilepsy</th>
<th>Stroke</th>
<th>Parkinson’s Disease</th>
<th>Migraine</th>
<th>Multiple sclerosis</th>
<th>Motor Neurone disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total N</td>
<td>453470 (91.20)</td>
<td>6758 (1.36)</td>
<td>12755 (2.67)</td>
<td>3315 (0.67)</td>
<td>18254 (3.67)</td>
<td>2315 (0.47)</td>
<td>559 (0.11)</td>
</tr>
<tr>
<td>Age, mean (SD), years</td>
<td>57.35 (8.09)</td>
<td>58.15 (8.11)</td>
<td>62.24 (6.66)</td>
<td>63.95 (5.29)</td>
<td>56.32 (7.95)</td>
<td>56.16 (7.66)</td>
<td>61.68 (6.64)</td>
</tr>
<tr>
<td>Sex</td>
<td>Female</td>
<td>245508 (54.14)</td>
<td>3397 (90.27)</td>
<td>5099 (39.98)</td>
<td>1265 (38.16)</td>
<td>14105 (77.27)</td>
<td>1679 (72.53)</td>
</tr>
<tr>
<td>Established diagnoses at study baseline (post-diagnosis)</td>
<td></td>
<td>4247 (0.85)</td>
<td>6352 (1.28)</td>
<td>873 (0.18)</td>
<td>14588 (2.93)</td>
<td>1861 (0.37)</td>
<td>70 (0.01)</td>
</tr>
<tr>
<td>Median time since diagnosis (to study baseline), years</td>
<td></td>
<td>23.74</td>
<td>5.59</td>
<td>4.15</td>
<td>12.61</td>
<td>12.46</td>
<td>3.28</td>
</tr>
<tr>
<td>New diagnoses after study baseline (pre-diagnosis)</td>
<td></td>
<td>2511 (0.51)</td>
<td>6403 (1.29)</td>
<td>2442 (0.49)</td>
<td>3666 (0.74)</td>
<td>454 (0.09)</td>
<td>489 (0.10)</td>
</tr>
<tr>
<td>Median time to diagnosis (after study baseline), years</td>
<td></td>
<td>-7.21</td>
<td>-7.59</td>
<td>-8.44</td>
<td>-7.62</td>
<td>-6.61</td>
<td>-6.78</td>
</tr>
<tr>
<td>Education</td>
<td>Higher</td>
<td>212761 (46.92)</td>
<td>2540 (37.59)</td>
<td>4543 (35.62)</td>
<td>1377 (41.54)</td>
<td>8876 (48.62)</td>
<td>1129 (48.77)</td>
</tr>
<tr>
<td></td>
<td>Upper Secondary</td>
<td>57133 (12.6)</td>
<td>858 (12.7)</td>
<td>1517 (11.89)</td>
<td>347 (10.47)</td>
<td>2211 (12.11)</td>
<td>225 (9.72)</td>
</tr>
<tr>
<td></td>
<td>Lower Secondary</td>
<td>24699 (5.45)</td>
<td>333 (4.93)</td>
<td>580 (4.55)</td>
<td>176 (5.31)</td>
<td>1043 (5.71)</td>
<td>154 (6.65)</td>
</tr>
<tr>
<td></td>
<td>Vocational</td>
<td>75246 (16.59)</td>
<td>1166 (17.25)</td>
<td>1963 (15.39)</td>
<td>484 (14.6)</td>
<td>3191 (17.48)</td>
<td>438 (18.92)</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>83631 (18.44)</td>
<td>1861 (27.54)</td>
<td>4152 (32.55)</td>
<td>931 (28.08)</td>
<td>2933 (16.07)</td>
<td>369 (15.94)</td>
</tr>
<tr>
<td>Socioeconomic status quintile</td>
<td>1 (least deprived)</td>
<td>91440 (20.16)</td>
<td>1041 (15.4)</td>
<td>2025 (15.88)</td>
<td>721 (21.75)</td>
<td>3572 (19.57)</td>
<td>467 (20.17)</td>
</tr>
<tr>
<td></td>
<td>2-4</td>
<td>272574 (60.11)</td>
<td>3777 (55.89)</td>
<td>7186 (56.34)</td>
<td>1933 (58.31)</td>
<td>10970 (60.1)</td>
<td>1391 (60.09)</td>
</tr>
<tr>
<td></td>
<td>5 (most deprived)</td>
<td>88891 (19.6)</td>
<td>1938 (28.68)</td>
<td>3529 (27.67)</td>
<td>656 (19.79)</td>
<td>3686 (20.19)</td>
<td>453 (19.57)</td>
</tr>
<tr>
<td></td>
<td>n/a</td>
<td>565 (0.12)</td>
<td>2 (0.03)</td>
<td>15 (0.12)</td>
<td>5 (0.15)</td>
<td>26 (0.14)</td>
<td>4 (0.17)</td>
</tr>
</tbody>
</table>

Percentages may not sum to 100 because of rounding.
Higher education defined as college/university degree or other professional qualification; upper secondary, second/final stage of secondary education; lower secondary, first stage of secondary education; vocational, work-related practical qualifications.

Socioeconomic status assessed on the Townsend deprivation index, which combines information on social class, employment, car availability, and housing.

Table 2. Number of participants in the imaging sub-cohort and temporal relation to diagnosis across different neurological conditions

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Controls N (%)</th>
<th>Epilepsy N (%)</th>
<th>Stroke N (%)</th>
<th>Parkinson's Disease N (%)</th>
<th>Migraine N (%)</th>
<th>Multiple sclerosis N (%)</th>
<th>Motor Neurone disease N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total N (imaging cohort)</td>
<td>42468 (93.57)</td>
<td>370 (0.87)</td>
<td>636 (1.50)</td>
<td>94 (0.22)</td>
<td>1973 (4.65)</td>
<td>170 (0.40)</td>
<td>19 (0.04)</td>
</tr>
<tr>
<td>Established diagnoses at brain scan visit (post-diagnosis)</td>
<td>-</td>
<td>321 (0.76)</td>
<td>468 (1.10)</td>
<td>56 (0.13)</td>
<td>1852 (4.36)</td>
<td>155 (0.36)</td>
<td>3 (0.07)</td>
</tr>
<tr>
<td>Median time since diagnosis (to brain scan), years</td>
<td>-</td>
<td>27.36</td>
<td>10.87</td>
<td>6.25</td>
<td>33.68</td>
<td>17.00</td>
<td>2.62</td>
</tr>
<tr>
<td>New diagnoses after brain scan visit (pre-diagnosis)</td>
<td>-</td>
<td>49 (0.12)</td>
<td>168 (0.40)</td>
<td>38 (0.09)</td>
<td>121 (0.28)</td>
<td>15 (0.04)</td>
<td>16 (0.04)</td>
</tr>
<tr>
<td>Median time to diagnosis (after brain scan), years</td>
<td>-</td>
<td>-1.80</td>
<td>-1.60</td>
<td>-2.26</td>
<td>-1.56</td>
<td>-1.39</td>
<td>-1.97</td>
</tr>
</tbody>
</table>

Percentages may not sum to 100 because of rounding.
References


24. Fawns-Ritchie C, Deary I. Reliability and validity of the UK Biobank cognitive tests. Published online 2019. doi:10.1101/19002204


34. Manohar SG. Matlab: MATLAB tools for plotting, data analysis, eye tracking and experiment design (Public). Published online 2019. doi:10.17605/OSF.IO/VMABG


