Machine learning spots the time to treat Huntington disease

Igor Koval
Paris Brain Institute
Thomas Dighiero-Brecht
Hôpital Pitié-Salpêtrière
Allan Tobin
Brain Research Institute UCLA
Sarah Tabrizi
University College London https://orcid.org/0000-0003-2716-2045
Rachael Scahill
UCL Institute of Neurology
Stanley Durrleman
Institut du Cerveau et de la Moelle Épinière
Alexandra Durr (alexandra.durr@icm-institute.org)
Hôpital Pitié-Salpêtrière https://orcid.org/0000-0002-8921-7104

Article

Keywords: Huntington disease, neurodegenerative diseases, preventive treatment

Posted Date: March 9th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-264531/v1

Abstract

We propose a new approach, based on machine learning, to evaluating new treatments for neurodegenerative diseases. Using data from two longitudinal studies of 299 participants with early Huntington Disease, we learned the range of likely trajectories of 15 imaging and clinical biomarkers from the premanifest to manifest disease stages. We positioned independent 11,510 patients on these maps using their baseline data and hence forecast the values of their biomarkers at any timepoint. Applied to trial design, we showed that sample size can be decreased by up to 50% by selecting individuals for whom we predict a significant change for their outcome measures during trial. Reduction occurs whatever the selected outcome measures and the targeted disease stage. This approach does not only select the right patient at the right time for the right trial, but also guides decisions about when to start preventive treatments.

Introduction

No treatments now exist to slow disease progression in any neurodegenerative disorder. So far, disease-modifying treatments have been uniformly unsuccessful, leaving patients, families and practitioners with limited therapeutic options.

The repeated failures of potentially preventive interventions underscore the concern that treatments given too late are not effective. While some drugs may interfere with the underlying pathological process, late intervention, so far, was not been enough to restore brain function. Nor is earlier treatment by itself an answer: treating apparently healthy people for decades raises both practical and ethical questions. An important step in therapeutic development, then, is finding the optimal time window for intervention, which may be narrower than expected.

Identifying the optimal moment for each person is a difficult challenge, however. Disease progression is complex, with changes revealed by clinical and imaging biomarkers. Furthermore, the trajectories of these changes are not linear, and their timing, ordering and pace are specific to each patient.

We show here that tailored machine learning tools can help clinicians and drug developers to anticipate disease progression and to identify the optimal moment for intervention. These tools use previously observed biomarker trajectories to forecast biomarker changes for new subjects.

Our approach can help clinical centres identify patients just before the onset of specific symptoms or of biomarker changes. This precise selection should allow a more effective evaluation of putative treatments. Once disease-modifying drugs are available, such a tool can also help ensure that each patient is treated at the right time, not too early neither too late, improving the benefit-to-risk ratio and limiting costs.

We have evaluated our approach with data from Huntington disease (HD), where we can unambiguously identify mutation carriers, based on genetic testing, who will develop clinical signs and symptoms. HD is

the most common of the polyglutamine diseases, all of which are caused by a pathological expansion of a CAG triplet repeat tract. In HD, expansion of the affected gene (*HTT*) above 36 repeats leads to disease, and the longer the repeat, the earlier the disease onset and the more rapid the progression.[1,2] Decades before disease onset, even at birth or prenatally, genetic testing can identify people who will develop the disease.

Despite the general trend of earlier onset with greater CAG expansions, prediction of disease onset for an individual at risk to develop HD is difficult: two individuals with the same length repeat may develop clinical disease decades apart.[3] The challenge for physicians and drug developers, as well as for eventual patients and families, is not only to predict the timing and dynamics of motor, behavioural and cognitive impairments, but also, crucially, to minimize the number of participants needed for a clinical trial.

Since 2004, several groups have developed methods to predict disease onset from pre-symptomatic phases.[4–7] These predictions have used several predictive indices, (1) burden or CAG-Age Product (CAP), (2) CAG-Age Product Scaled (CAPS), (3) the PIN index, which uses the values of motor and cognitive assessments in addition to the number of CAG repeats and age, and (4) the multivariate risk score base (MRS), which incorporates 34 additional variables.[4–7] Several studies have also used brain imaging or CSF neurofilament proteins to stratify premanifest individuals.[8–10]

All these indices improve predictions of HD onset, but so far none has been used to stratify participants for clinical trials. Still, Paulsen and colleagues showed that they could identify populations most likely develop the disease within three years, using Multivariate Risk Score (MRS), and to a lesser extent the PIN.[11] Further, Langbehn and colleagues show that the PIN index can increase the effect size of hypothetical treatments for a range of possible clinical endpoints in addition to disease onset.[12] This property of the PIN index is only incidental, however, since PIN has been computed to predict clinical onset, not the progression of particular biomarkers. A PIN adapted to each type of biomarker would most likely give better results, but it is not possible to create and validate as many indexes as biomarkers and to adjust threshold values according to a targeted therapeutic window.

We have therefore taken a different approach, using longitudinal observational data to describe the progression of 15 biomarkers. By combining data from patients at different stages, we can recreate trajectories of progression over long periods of time.[13,14] Once trained, our program can automatically incorporate data for a new subject. From the values of the biomarkers of a new participant at a given time, we can now predict future biomarker trajectories for that same mutation carrier. By using a single algorithm, we can thus anticipate changes of any biomarker at any stage of the disease. With our ability to predict biomarker trajectory and clinical signs, we can reduce the sample size for a clinical trial, whatever selected outcome measures and the targeted disease stage. Our strategy is systematically better than those based on CAP or PIN indexes. Our method will allow clinicians to improve patient monitoring and, once treatments are available, to make informed decisions about the timing of treatments.

Results

HD COURSE MAP depicts the progression of 15 radiological and clinical outcomes across all disease stages

We established HD COURSE MAP with longitudinal data from TRACK-HD and TRACK-ON, studies with 299 participants followed for 7 years (1,333 visits).
Subjects included both asymptomatic and symptomatic carriers of the pathological expansion (Table 1).
Figure 1, panels a, b, and c, illustrates the method we used to build HD COURSE MAP from these data.

Table 1
Demographic description of participants from the two TRACK studies (TRACKHD and TRACK ON) and the ENROLL-HD study.

Variable description and characteristics	TRACK HD & TRACK ON	ENROLL
Number of participants (and of visits)	299 (1,333)	11,510 (28,189)
Distributions of follow-up visits	39 participants with 2 visits, 46 with 3, 118 with 4, 11 with 5, 5 with 6 and 80 with 7	3,599 participants with 1 visit, 2,764 with 2, 2,535 with 3, 1,777 with 4, 665 with 5, 166 with 6 and 4 with 7
Age (± std) years old	45.0 (± 10.2)	49.0 (± 13.8)
Number of CAG repetitions (± std)	43.2 (± 2.7)	43.6 (± 3.6)
Woman/ Men / Unknown (%)	53.2 / 45.5 / 1.3	53.9 / 46.1 / 0
Education level (± std)	Number of years (± std): 3.8 (± 1.2)	ISCED level (± std): 3.6 (± 1.2)
Clinical features	Total motor score from the Unified Huntington disease Rating Scale (TMS) Total Functional Capacity from the UHDRS (TFC), and the apathy rating from the Problem Behavior Assessment (PBA-Apathy)	Total motor score from the Unified Huntington disease Rating Scale (TMS) Total Functional Capacity from the UHDRS (TFC), and the apathy rating from the Problem Behavior Assessment (PBA-Apathy)
Cognitive features	Stroop Word Reading Test (Stroop), Symbol Digit Modality Test (SDMT), Direct Circle Tracing (Circle Tracing), and Indirect Circle Tracing (Circle Tracing Indirect).	Stroop Word Reading Test (Stroop), and Symbol Digit Modality Test (SDMT)

Variable description and characteristics	TRACK HD & TRACK ON	ENROLL
Imaging features (volumes)	1. Striatum, 2. Globus Pallidus, 3. Putamen, 4. Caudate, 5. White matter, 6. Grey matter, Ventricles and total brain	Not done

The map is a long-term trajectory constructed from these 299 participants at different disease stages. It depicts the course of 15 radiological and clinical biomarkers along a normalized time axis, which we call “Huntingtonian Age “(HA). Participants had HAs between 30 and 80 years, so our model represents about 50 years of the natural history of the disease.

Figure 2a shows the progression of six of these biomarkers, two clinical evaluations (total motor score [TMS] and total functional score [TFC], two imaging markers (striatal and grey matter volumes), and two cognitive markers (Stroop Word Reading Test and Symbol Digit Modality Test [SDMT]). Some features, such as TMS, show sharp increases at some point. Others, including Stroop and striatal volume, progress in a nearly linear fashion. (See Supplementary Figure S1 for the progression of all 15 biomarkers)

Figure 2b and 2c show the distinct trajectories of these six biomarkers in two subgroups: subjects with fewer than 42 CAG repeats, and those with more than 44 repeats. As expected, subjects in the group with the longest repeat lengths had earlier onset and faster progression.

The 15 markers deviate from control values following the sequence shown in Fig. 3. (Supplementary Figure S3 shows details of the method). The first biomarker to diverge from controls who do not have the HD-causing allele is TMS, at HA of 37 years and an interquartile confidence interval [CI] from 34.5 to 39.2 years. TMS is the most powerful biomarker of early stages. Generally, both subjects and doctors become aware of onset only well after the first motor changes.

Changes of basal ganglia volumes are next: striatum at 42.2 years (CI=[39.8,43.4]); putamen at 42.9 years (CI=[40.2,44.3]); globus pallidus at 43.0 years (CI=[40.7,45.1]); and caudate at 46.8 years (CI=[44.1,48.3]). These volumetric imaging markers are well suited for studies of early and even premanifest individuals.

On the other hand, many biomarkers change only 20 years after motor onset. These late biomarkers include most clinical measurements as well as volume changes of whole brain, ventricles, total white matter, and total grey matter. These latter markers are therefore ill suited for studies of early stages.

TFC decline is a feature only of advanced HD. Despite the regulatory appeal of a functional readout, our analysis shows that TFC is an inappropriate endpoint for early-stage trials.

TRACK-HD and TRACK-ON did not measure neurofilament light chain in CSF, which may represent an important marker of early changes.[10]

Overall, HD COURSE MAP succinctly summarizes the natural history of HD biomarkers. Our conclusions are consistent with current knowledge of the disease.

HD COURSE MAP summarizes the variability in disease progression across subjects.

HD COURSE MAP not only represents the general trajectory of disease progression, but it also summarizes the trajectories of individual readouts in each participant. For each subject, three parameters describe these changes: (1) time-shift [τ], (2) acceleration [α], and (3) 15 inter-marker spacings [ω] (Fig. 1c). The first two parameters (τ and α) summarize differences in the timing and rate of progression. They map the actual age of the participant to HA.

The intermarker spacing parameters (ω_i) represent the different trajectories of individual biomarkers, accounting for individual phenotypic differences at a given HA. HD COURSE MAP summarizes the range of biomarker trajectories and captures the variability of progression among mutation carriers.

Together with the general trajectory determined by our analysis, these parameters allowed us to describe the progression of each biomarker for each participant. By comparing these parameters across subjects, we examined the influence of socio-demographics and genetics (Supplementary Figure S4). Our analysis showed that some cofactors (such as longer CAG repeats) modulate HD progression, while others (such as educational level or sex) do not. On average, the disease starts 2.4 years earlier for every CAG repetition (p-value: $5.0\ 10^{-25}$, CI=[2.756, 1.980]), and the rate of progression is multiplied by 1.043 for every CAG repetition (p-value = $4.3\ 10^{-4}$, CI=[1.020, 1.067]). Increased CAG repeats also advance degradation of Stroop test (p = $5.5\ 10^{-3}$) and delays grey-matter changes (p = $5.4\ 10^{-8}$), brain atrophy (p = $4.9\ 10^{-4}$) and ventricles volumes (p = $5.1\ 10^{-4}$) (Supplementary Figure S4).

HD COURSE MAP forecasts individual biomarkers up to four years in advance.

Using only baseline data, we can place each new mutation carrier on the HD COURSE MAP, estimating the parameters that allow the best fit to the general curve. We then used these parameters to predict each subject's progression curves for each biomarker. These curves forecast the value of all the biomarkers. We compare then these predictions with data at follow-up visits (Fig. 1d).

Baseline data are sufficient to determine inter-marker spacings (ω_i), the relative timing of biomarker changes. In principle, we would need data from two time-points to estimate the dynamic parameters α and τ. In practice, however, baseline data are often sufficient if we restrict the dynamic parameters to the range of values already observed in the training data.

We evaluated this prediction in two independent groups: (1) the TRACK/TRACK-ON subjects and (2) an independent set of HD subjects ("ENROLL").

We first undertook a cross-validation with the TRACK HD/ON subjects. Using only 80% of the subjects, we redetermined the generalized HD COURSE MAP. We then predicted biomarker progressions for the remaining 20%. We compared the predicted values to the data at follow-up visits one, two, three and four years later. Figure 4 shows the distribution of prediction errors. For instance, once rescaled on a 0-100 scale, the mean absolute error on TMS is of 2.97 at one year, 3.93 at two years, 4.55 at three years and 2.77 at four years which is of the order of the error in the measurements.[15]

We repeated this process with 11,510 ENROLL participants. We positioned each patient on the HD COURSE MAP, using only baseline data. We then evaluated the distribution of prediction errors after one, two, three, and four years (Fig. 4). For example, the mean absolute errors for ENROLL subjects are respectively of 5.04, 5.77, 6.43 and 6.84 for the TMS. Although HD COURSE MAP used data from TRACK-HD and TRACK-ON only, we could predict biomarker progression of subjects from the ENROLL registry with a mean absolute error of less than 5% for all biomarkers.

We have also applied our method to increasing cognitive impairment in Alzheimer's disease. Our predictions were better than the 56 alternative methods in an open data challenge.[16,17] We know of no alternative method that can predict the values of a range of HD biomarkers.

HD COURSE MAP can select participants for clinical trials before significant disease progression.

Our ability to predict the trajectories of outcome measures for each mutation carrier can greatly improve subject selection for clinical trials. Knowledge of the temporal ordering of biomarker allows us to select outcome measures that progresse most rapidly at the stage targeted by the drug (Fig. 3). Trialists will then be able to select participants close to the time when selected outcome measures will change. This approach will reduce the fraction of subjects in whom an outcome measure is unlikely to change during the course of a trial.

To test the ability of HD COURSE MAP to help clinical trial design, we simulated four clinical trials, each centred on a different stage of disease development, from pre-manifest to early symptomatic (Table 2). We determined disease stage using TMS and TFC, with TMS < 15 as premanifest or early HD. The maximum, or worst, TMS score is 124, and HD-mutation carriers with normal motor examination have a TMS < 5.[18] Subjects with TMS between 5 and 15 have only early signs, and they may not recognize these signs as pathological. After choosing the disease stage and the corresponding best-suited endpoint for

each trial, we further selected subjects and markers for whom HD COURSE MAP predicted a significant change during trial duration.

We compared our simulations with trials with unselected participants or with two alternative selection criteria, based on Burden Score or PIN score (Fig. 5). We then computed, for each selection, the sample size needed to detect hypothetical treatment effects (Fig. 6).

Table 2

Trial designs to compare selection procedures and sample sizes (Fig. 6). Simulations target different disease stages, trial durations and outcome measures. Three designs aim to reproduce the ones of the Crest-D, Pride-HD and 2Care-HD trials.[26–28]

	Premanifest or very early stages defined by motor score (maximum worse value TMS = 124)			**Early affected based on functional (maximum worst value TFC = 13)**
Inclusion criteria	$0 \leq TMS \leq 10$	$0 \leq TMS \leq 5$	$5 \leq TMS \leq 15$	$7 \leq TFC \leq 10$
Trial duration	2 years	4 years	3 years	3 years
Outcome measure	Rate of change of normalized caudate volume	Change of absolute TMS	Rate of change of TMS	Rate of change of TMS
Treatment effect	15–30% less change in the caudate rate	1.5 to 3.0 points less change in the TMS	25–40% less in the TMS rate of change	30–50% less in the TMS rate of change
Experiments	319 patients from TRACK HD / ON (using cross-validation)	HD Course map estimated on TRACK-HD/ON. 330 predictions on ENROLL.	HD Course map estimated on TRACK-HD/ON. 665 predictions on ENROLL.	HD Course map estimated on TRACK-HD/ON. 527 predictions on ENROLL.

Our first simulation started with 527 ENROLL participants in early symptomatic stages, with a baseline functional score (TFC) between 7 and 10.

We simulated a trial whose endpoint was a 40% reduction in the rate of TMS change over the 3 years of the study. All three selection methods decreased the number of required participants. HD COURSE MAP was the most effective, requiring a sample size of 150, while the Burden Score required nearly 300, and PIN score 220, compared to the 527 initially selected. HD COURSE MAP therefore reduced the number of participants by 50% compared to the burden score and 32% compared to the PIN index.

The second simulation, which sought a 30% reduction after 3 years in the rate of TMS change, began with 630 ENROLL participants with TMS between 5 and 15. In this case, PIN and HD COURSE MAP had similar performance, yielding a sample size of 360 while Burden Score required a sample size of 400. In this case, HD COURSE MAP reduced the number of needed inclusions by 10% compared to the Burden Score.

The third simulation started with 326 ENROLL participants who were *asymptomatic*, with TMS < 5. The endpoint of this simulated trial was a 3-point decrease in TMS at the end of 4 year trial. PIN and Burden Score identified participants whose progressions were not particularly different from the starting cohort, leading to a greater sample size than initially.

By contrast, HD COURSE MAP required a sample size of 85 participants, as opposed to 100 without any selection, yielding a 15% reduction.
Finally, we simulated a trial in very early stage mutation carriers (TMS < 10) using only imaging as an endpoint, rather than clinical evaluation. The trial was designed to detect a 25% decline in the rate of caudate atrophy in two years. Because ENROLL-HD did not include imaging data, we used 108 TRACK HD/ON participants. Again, HD COURSE MAP allowed the best reduction of sample size, with a required sample size of 210, compared to 400 for burden score and 320 for PIN.

Here our method reduced the number of needed participants by 48% and 34%.
Our proposed selection method demonstrates a significant benefit at any disease stage. For presymptomatic mutation carriers—for whom therapeutic intervention could prove the most beneficial—only HD COURSE MAP reliably identifies participants likely to be informative in a clinical trial.

Discussion

We present an interpretable machine learning model, HD COURSE MAP, that allowed us to position each HD mutation carrier onto an idealized timeline and predicted the values of their outcome measures up to four years.

Using this map to identify HD individuals likely to progress in the course of a trial, we allow a significant reduction (up to 50%) in the number of participants necessary for each clinical trial when compared to current selection methods.
This method is applicable for a wide range of disease stages and outcome measures, even for premanifest stages. In all cases, our method will help trial designers to maximize their chances of success. HD is a rare disease, conducting trials with fewer participants while maintaining statistical power is particularly important.

With the development of new drugs, treatment of neurodegenerative diseases will require personalised therapeutic strategies, or precision medicine. This approach has already benefited cancer patients, whose treatments differ with the genetic characteristics of their cancers. For neurodegenerative diseases, a

critical question for precision medicine will be the optimal time to treat. Our data-driven approach will improve the ability to choose that time and will provide an important tool for future care.

Previous approaches to these questions for HD have been neither versatile nor easily explained. Our approach, HD COURSE MAP, starts by building a "digital twin" for each patient, based on a set of biomarkers determined at a single time. Our program allows both physician and patient to see, understand, and anticipate disease progression several years in advance.

HD COURSE MAP can enhance not only the efficiency of clinical trials but also the ability to follow disease progression and to prescribe effective treatments.

Methods

Participants

The study used data from two multicenter observational studies: TRACK-HD and TRACK-ON.[18–21] These studies measured motor, cognitive, and neuropsychological biomarkers. They also determined the volumes of several brain structures, using methods developed at the University of Iowa and at University College London.[22,23] The purpose of these studies was to identify features sensitive to disease progression, starting well before anticipated clinical onset. TRACK-HD followed 117 pre-symptomatic HD-mutation carriers and 116 patients with early-stage HD. TRACK-ON's 239 participants included previous TRACK-HD participants as well as new presymptomatic HD-mutation carriers. These studies also included 186 "control" subjects, most of whom were sibs or other family members without a pathogenic HD-mutation.

In order to validate our prediction method in an independent cohort, we used data for ENROLL-HD, a global clinical research platform designed to facilitate clinical research in HD (https://enroll-hd.org).

As part of this multicenter, longitudinal observational study, core data sets are collected annually from all research participants.
Using a risk-based monitoring approach, all collected data are monitored for accuracy and quality.

Overall, we used data from 11,510 participants with a total number of visits of 28,189.

Feature selection and preprocessing

We selected 15 features in our initial modelling of biomarkers' progression based on TRACK-HD and TRACK-ON participants:

- Clinical (1) TMS, the total motor score from the Unified Huntington disease Rating Scale [UHDRS], (2) Total Functional Capacity [TFC] from the UHDRS and (3) the apathy rating from the Problem Behavior Assessment [PBA-Apathy].
- Cognitive (1) Stroop Word Reading Test [Stroop], (2) Symbol Digit Modality Test [SDMT], (3) Direct Circle Tracing, and (4) Indirect Circle Tracing.

- Imaging (1) striatum, (2) globus pallidus, (3) putamen, (4) ventricles, (5) caudate, (6) white matter, (7) grey matter, and (8) total brain volumes– all normalized to total intracranial volume.

Each of these measurements increased or decreased monotonically. To deal with all 15 biomarkers in the same way, we adopted the convention that their values increase with progression. For cases where measures decrease with progression (such as caudate volume and SDMT) we reverse the temporal profile. We normalize each reading to a scale of zero to + 1, taking the value for controls as 0 and the value of the pathological state as + 1. This operation is easy for values that have inherent maximum and minimum values, such as TMS, SDMT, TFC and PBA Apathy.

We adopted a different normalization for the unbounded features such as Circle Tracing and Stroop tests, both measured in seconds, and imaging-derived brain volumes. For the 0 of each feature, we selected the 95th percentile of the control distribution. For the + 1 of the scale, we take the maximum value of each feature among all the subjects.

In each case, most temporal profiles start at values near 0 (control value), increasing to 1 (abnormal values) with disease progression. Some features, such as volume and SDMT start with values greater than zero, even for controls.

Multimodal digital model of disease progression (HD COURSE MAP)

HD COURSE MAP is built on a new statistical learning technique named disease course mapping.[13,14,16,24] The method has two goals: (1) to model the long-term changes of clinical and imaging biomarkers from the repeated observation of multiple subjects at various disease stages; (2) to predict the future values of these biomarkers from measurements at the entry of each new participant.

The approach resembles that of the Talairach brain atlas, which maps brain structures independently of individual variations in the size and overall shape of the brain.[25] In this atlas, each structure has its unique set of (x,y,z) coordinates, allowing it to be located easily and reproducibly. Additionally, images from any new subject can be transformed to be superimposed on the atlas. In the same manner, the method places subjects along a time axis representing the stage of disease progression. For a single biomarker, each value along an idealized logistic progression indicates a moment on the time axis that we call “Huntingtonian age (HA).” While HA increases with disease progression, it does not directly correspond to the subject’s Actual Age (AA).

Using data from three hypothetical subjects, we sketch two idealized curves corresponding to the progression of two biomarkers (Fig. 1). We can fit individual data onto the two idealized curves of Fig. 1B using three computed parameters:

(1) time-shift τ_i, the temporal onset of disease progression for all features in subject i; τ_i is negative for earlier-than-average progression and positive for later-than-average progression,

(2) acceleration $\boldsymbol{\alpha}_i$, the change in the rate of progression of subject i; $\boldsymbol{\alpha}$ =1 means that, for that subject, all the biomarker trajectories match the idealized curves after adjusting the timing with $\boldsymbol{\tau}$ and $\boldsymbol{\omega}$. When $\boldsymbol{\alpha}$ > 1, it takes less time for the same changes to occur, e.g. $\boldsymbol{\alpha}$=2 means that it will twice as fast as average for that subject to undergo a given change in a biomarker trajectory,

(3) intermarker spacing $\boldsymbol{\omega}_{ij}$, the temporal ordering of biomarker j with respect to other biomarkers for subject i. The sign of $\boldsymbol{\omega}_{ij}$, indicates whether a single feature in a particular subject exhibits earlier or later changes relative to other features (after normalizing the age at onset $\boldsymbol{\tau}$ and pace of progression $\boldsymbol{\alpha}$): $\boldsymbol{\omega}_{ij}$ < 0 means earlier degradation, and $\boldsymbol{\omega}_{ij}$ > 0 means later.

The first two parameters allow us to map the actual age of the participant onto his Huntingtonian Age (HA) using the affine transformation $HA_{ij} = \boldsymbol{\alpha}i *(t_{ij} - \boldsymbol{\tau}_i)$, where t_{ij} is the age of the i-th subject at her/his j-th visit, and HA_{ij} his corresponding Huntingtonian Age. This computation resembles the burden score except that it includes both a multiplicative and an additive factor. We also treat, as separate parameters, the age at onset and the pace of progression. Further, the pace of progression is estimated from the participant's biomarkers and is not fully determined by the number of CAG repeats. We verified *a posteriori* that these parameters are both indeed dependent on the number of CAG repeats.

The idealized set of logistics thus serves as a reference pattern of biomarker development. Our fitting to this set of curves requires simultaneous estimation of the individual parameters ($\boldsymbol{\tau}_i$, $\boldsymbol{\alpha}_i$, and $\boldsymbol{\omega}_{ij}$) of each individual in a longitudinal setting. Each mutation carrier has a single $\boldsymbol{\tau}$, a single $\boldsymbol{\alpha}$, and one $\boldsymbol{\omega}$ for each biomarker (15 in this analysis). We determined the shape and position of the logistic curves so that mean value of $\boldsymbol{\tau}_i$ across all subjects is 0, the mean value of the logarithm of $\boldsymbol{\alpha}_i$ is 0, and the mean value of each $\boldsymbol{\omega}_{ij}$ is 0. The resulting idealized set of logistic curves represents therefore the typical subject in the studied population. All in one, the estimation of the disease course map takes the form a non-linear Bayesian mixed-effects model where the shape and position of the logistics curves represent the fixed effects and the subject's parameters ($\boldsymbol{\tau}_i$, $\boldsymbol{\alpha}_i$, and $\boldsymbol{\omega}_{ij}$) the random effects. We iteratively performed joint estimations of these population parameters with the individual ones to minimize errors between the idealized curves and the individual projections.[11] The result is called HD COURSE MAP.

We can apply the reference model to baseline data of a new participant. We estimate the individual parameters for the subject and then transform the idealized logistic curves to match empirical data (Fig. 1D). This process accomplishes two goals: (1) characterization of subjects according to values of $\boldsymbol{\tau}$, $\boldsymbol{\alpha}$ and $\boldsymbol{\omega}$ (Figs. 1B and 1C); and (2) prediction of future biomarker values from our data-derived curves (Fig. 1D).

We use the open-source software Leaspy to estimate the model parameters from a longitudinal data set and fit the trained model to new data. The software is publicly available at: https://gitlab.com/icm-institute/aramislab/leaspy/

Experimental design

The estimation of HD COURSE MAP on TRACK-HD and TRACK-ON data required a minimum of two visits per participant. Our analysis therefore included 299 participants, and 1,333 visits (for further demographics, see Table 1). To estimate confidence intervals, we performed 100 bootstrap runs. We compared the mean progression of each feature in terms of HA to predefined thresholds, thereby revealing the temporal ordering of the deviation of the biomarkers from their control values (Supplementary Figure S3).

Our analysis allowed us to predict the values of individual features up to four years after the first evaluation of each participant, that is, without the need of longitudinal data.
To cross-validate our prediction method of TRACK participants, we examined its predictive power for 20% with eight features, using the other 80% for calibration.
This operation was repeated five times so that each participant was examined once. Further validation was performed on ENROLL-HD participants whose data were not used in the initial training thus representing an independent cohort for replication. We report the error between the predicted values and the actual values.

We finally demonstrate the value of our predictions by applying it to identify mutation carriers most likely to reveal a drug effect during clinical trials targeting early HD patients. To do so, we simulate multiple clinical trials that differ in term of durations, drug effects, inclusion criteria, endpoints and disease stage – see Table 2. Trial simulation was inspired from the design of past clinical trials, in particular Crest-D, Pride-HD and 2Care-HD.[26–28]

Declarations

ACKNOWLEDGMENTS

We are very grateful to the participants of TRACK-HD, TRACK-ON and ENROLL-HD, whose data were the basis of this analysis.

The work has been funded in part by the European Research Council (ERC) under grant agreement No 678304, European Union's Horizon 2020 research and innovation programme under grant agreement No 666992 (EuroPOND) and No 826421 (TVB-Cloud), and the program "Investissements d'avenir" ANR-10-IAIHU-06 (IHU ICM) and ANR-19-P3IA-0001 (PRAIRIE 3IA Institute).

The TRACK-HD and Track-On HD group members are Tanka Acharya, MSc, University of Iowa, Iowa City; Sophie Andrews, PhD, Monash University, Melbourne, Victoria, Australia; Natalie Arran, DClinPsy, University of Manchester, Manchester, United Kingdom; Eric Axelson, PhD, University of Iowa, Iowa City; Eric Bardinet, PhD, Institut du Cerveau et de la Moelle Epinière and Pitié-Salpêtrière University Hospital, Paris, France; Natalie Bechtel, MD, University of Münster, Münster, Germany; Claire Berna, MSc, University College London, London, United Kingdom; Stefan Bohlen, PhD, George-Huntington-Institute, Münster, Germany; Beth Borowsky, CHDI, New York, USA; Jenny Callaghan, BSc, University of Manchester, Manchester, United Kingdom; Amy Cassidy, MSc, London School of Hygiene & Tropical Medicine, London,

United Kingdom; Allison Coleman, BSc, University of British Columbia, Vancouver, British Columbia, Canada; David Craufurd, PhD, University of Manchester, Manchester, United Kingdom; Helen Crawford, PhD, University College London, London, United Kingdom; Rachelle Dar Santos, BSc, University of British Columbia, Vancouver, British Columbia, Canada; Joji Decolongon, MSc, University of British Columbia, Vancouver, British Columbia, Canada; Eve Dumas, PhD, Leiden University Medical Center, Leiden, the Netherlands; Mannie Fan, PhD, University of British Columbia, Vancouver, British Columbia, Canada; Chris Frost, MSc, London School of Hygiene & Tropical Medicine, London, United Kingdom; Rhia Ghosh, PhD, University College London, London, United Kingdom; Claire Gibbard, PhD, University College London, London, United Kingdom; Sarah Gregory, PhD, University College London, London, United Kingdom; Davina Hensman-Moss, PhD, University College London, London, United Kingdom; Nicola Hobbs, PhD, University College London, London, United Kingdom; Celine Jauffret, MSc, Institut du Cerveau et de la Moelle Epinière and Pitié-Salpêtrière University Hospital, Paris, France; Eileanoir Johnson, PhD, University College London, London, United Kingdom; Hans Johnson, PhD, University of Iowa, Iowa City; Rebecca Jones, MSc, London School of Hygiene & Tropical Medicine, London, United Kingdom; Caroline Jurgens, PhD, Leiden University Medical Center, Leiden, the Netherlands; Damian Justo, PhD, Institut du Cerveau et de la Moelle Epinière and Pitié-Salpêtrière University Hospital, Paris, France; Ruth Keogh, PhD, London School of Hygiene & Tropical Medicine, London, United Kingdom; Tamara Koren, PhD, University of British Columbia, Vancouver, British Columbia, Canada; Izelle Labuschagne, PhD, Monash University, Melbourne, Australia; Nayana Lahiri, MD, University College London, London, United Kingdom; Stephane Lehericy, PhD, Institut du Cerveau et de la Moelle Epinière and Pitié-Salpêtrière University Hospital, Paris, France; Bernhard Landwehrmeyer PhD, Ulm University, Ulm, Germany; Douglas Langbehn, PhD, University of Iowa, Iowa City; Blair Leavitt, PhD, University of British Columbia, Vancouver, British Columbia, Canada; Jeff Long, PhD, University of Iowa, Iowa City; Ian Malone, PhD, University College London, London, United Kingdom; Cecilia Marelli, MD, Institut du Cerveau et de la Moelle Epinière and Pitié-Salpêtrière University Hospital, Paris, France; Peter McColgan, PhD, University College London, London, United Kingdom; James Mills, PhD, University of Iowa, Iowa City; Kevin Nigaud, PhD, Institut du Cerveau et de la Moelle Epinière and Pitié-Salpêtrière University Hospital, Paris, France; Alison O'Regan, MSc, Monash University, Melbourne, Australia; Gail Owen, PhD, University College London, London, United Kingdom; Marina Papoutsi, PhD, University College London, London, United Kingdom; Tracey Pepple, BSc, University College London, London, United Kingdom; Terri Petkau, PhD, University of British Columbia, Vancouver, British Columbia, Canada; Sarah Queller, MA, Indiana University, Indianapolis; Joy Read, MSc, University College London, London, United Kingdom; Ralf Reilmann, PhD, University of Münster, Münster, Germany; Raymund Roos, PhD, Leiden University Medical Center, Leiden, the Netherlands; Miranda Say, MSc, University College London, London, United Kingdom; Anne Schoonderbeek, PhD, Leiden University Medical Center, Leiden, the Netherlands; Cheryl Stopford, PhD, University of Manchester, Manchester, United Kingdom; Julie Stout, PhD, Monash University, Melbourne, Victoria, Australia; Aaron Sturrock, MD, University of British Columbia, Vancouver, British Columbia, Canada; Ellen 't Hart, PhD, Leiden University Medical Center, Leiden, the Netherlands; Romain Valabrègue, PhD, Institut du Cerveau et de la Moelle Epinière and Pitié-Salpêtrière University Hospital, Paris, France; Simon van den Bogaard, MD, Leiden University Medical Center, Leiden, the Netherlands; Jeroen van der Grond, PhD, Leiden University Medical

Center, Leiden, the Netherlands; Chiachi Wang, PhD, University of Iowa, Iowa City; Nathalia Weber, PhD, University of Münster, Münster, Germany; Daisy Whitehead, MSc, University College London, London, United Kingdom; and Marie-Noelle Witjes-Ane, PhD, Leiden University Medical Center, Leiden, the Netherlands.

We are grateful to the CHDI foundation that has initiated and supported Track-HD.

AUTHOR CONTRIBUTIONS

IK: data processing, data analysis, data interpretation, writing. TDB: data processing, data analysis, data interpretation, writing. AJT: data interpretation, writing. SJT: data collection, writing. RIS: data collection, data processing, writing. SD: study design, data analysis, data interpretation, writing. AD: data collection, study design, data interpretation, writing

COMPETING INTERESTS

A patent has been filed by INSERM Transfer under the reference PCT/IB2016/052699 (inventors: J.-B. Schiratti, S. Allassonnière, O. Colliot, S. Durrleman). It applies to the use of disease course mapping for selecting patients into clinical trials. The patent has been accepted in the USA and is under investigation in Europe and Japan. SD received a Sanofi iDEA award from Sanofi for a collaborative research project. All other authors declare no competing interests.

References

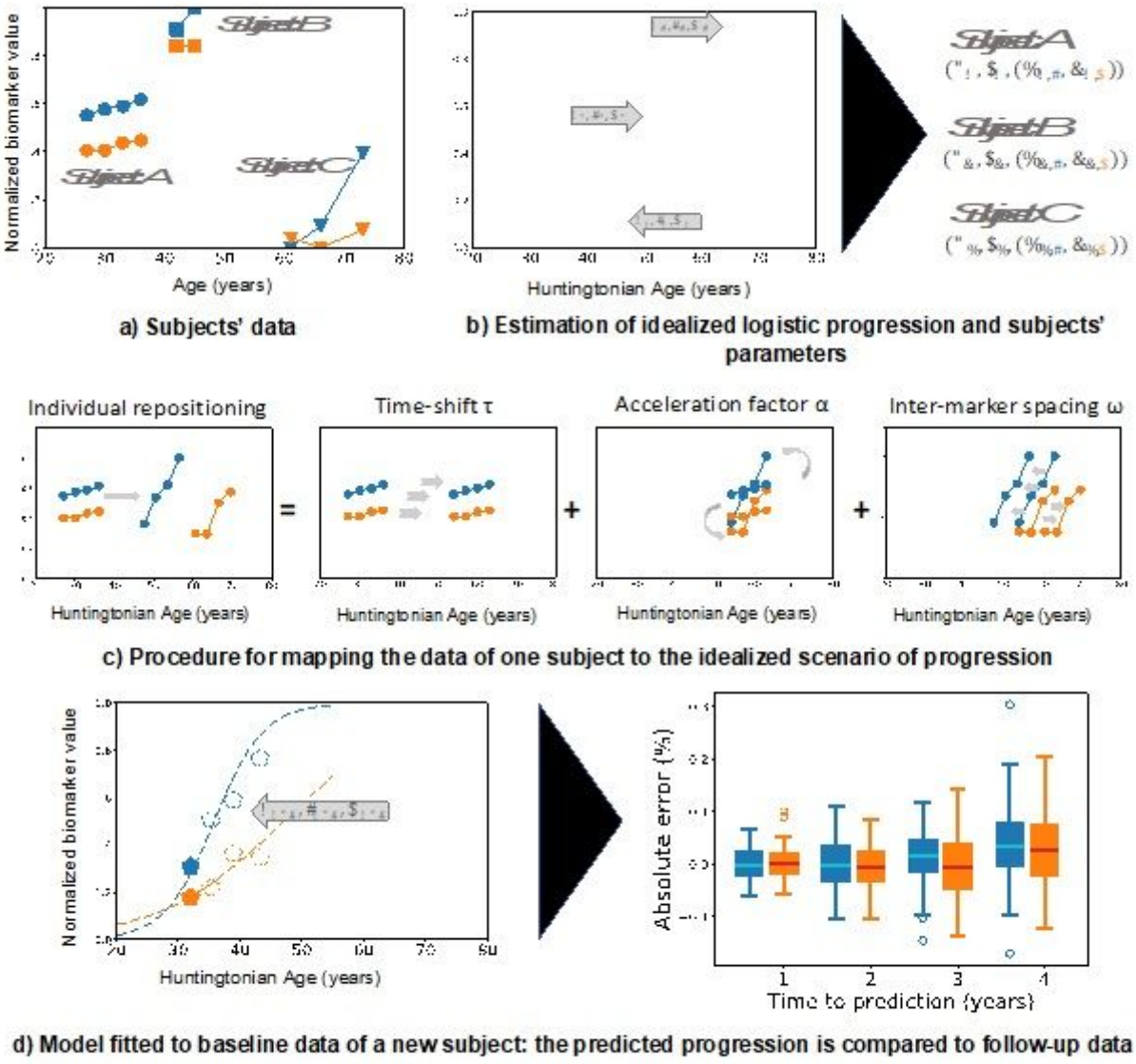
1. Tabrizi, S. J., Flower, M. D., Ross, C. A. & Wild, E. J. Huntington disease: new insights into molecular pathogenesis and therapeutic opportunities. *Nat. Rev. Neurol.* **16**, 529–546 (2020).
2. Andrew, S. E. *et al.* The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. *Nat. Genet.* **4**, 398–403 (1993).
3. Lee, J.-M. *et al.* CAG repeat expansion in Huntington disease determines age at onset in a fully dominant fashion. *Neurology* **78**, 690–695 (2012).
4. Langbehn, D. *et al.* A new model for prediction of the age of onset and penetrance for Huntington's disease based on CAG length: Prediction of the age of onset and penetrance for HD. *Clin. Genet.* **65**, 267–277 (2004).
5. Zhang, Y. *et al.* Indexing disease progression at study entry with individuals at-risk for Huntington disease. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **156**, 751–763 (2011).
6. Long, J. D., Paulsen, J. S., & PREDICT-HD Investigators and Coordinators of the Huntington Study Group. Multivariate prediction of motor diagnosis in Huntington's disease: 12 years of PREDICT-HD: Multivariate Prediction of HD Motor Diagnosis. *Mov. Disord.* **30**, 1664–1672 (2015).
7. Long, J. D. *et al.* Validation of a prognostic index for Huntington's disease: Prognostic Index For Huntington's Disease. *Mov. Disord.* **32**, 256–263 (2017).

8. Aylward, E. H. *et al.* Striatal Volume Contributes to the Prediction of Onset of Huntington Disease in Incident Cases. *Biol. Psychiatry* **71**, 822–828 (2012).
9. Polosecki, P. *et al.* Resting-state connectivity stratifies premanifest Huntington's disease by longitudinal cognitive decline rate. *Sci. Rep.* **10**, 1252 (2020).
10. Scahill, R. I. *et al.* Biological and clinical characteristics of gene carriers far from predicted onset in the Huntington's disease Young Adult Study (HD-YAS): a cross-sectional analysis. *Lancet Neurol.* **19**, 502–512 (2020).
11. Paulsen, J. S., Lourens, S., Kieburtz, K. & Zhang, Y. Sample enrichment for clinical trials to show delay of onset in huntington disease. *Mov. Disord.* **34**, 274–280 (2019).
12. Langbehn, D. R. & Hersch, S. Clinical Outcomes and Selection Criteria for Prodromal Huntington's Disease Trials. *Mov. Disord.* **35**, 2193-2200 (2020).
13. Schiratti, J.-B., Allassonnière, S., Colliot, O. & Durrleman, S. A Bayesian Mixed-Effects Model to Learn Trajectories of Changes from Repeated Manifold-Valued Observations. *J. Mach. Learn. Res.* **18**, 1-33 (2017).
14. Koval, I. *et al.* Spatiotemporal Propagation of the Cortical Atrophy: Population and Individual Patterns. *Front. Neurol.* **9**, 235 (2018).
15. Boringa, J. B. *et al.* The Brief Repeatable Battery of Neuropsychological Tests: normative values allow application in multiple sclerosis clinical practice. *Mult. Scler. J.* **7**, 263–267 (2001).
16. Durrleman, S. *et al.* AD Course Map charts Alzheimer's disease progression. 22.
17. Marinescu, R. V. *et al.* Predicting Alzheimer's disease progression: Results from the TADPOLE Challenge: Neuroimaging: Neuroimaging predictors of cognitive decline. *Alzheimers Dement.* **16**, (2020).
18. Tabrizi, S. J. *et al.* Biological and clinical manifestations of Huntington's disease in the longitudinal TRACK-HD study: cross-sectional analysis of baseline data. *Lancet Neurol.* **8**, 791–801 (2009).
19. Tabrizi, S. J. *et al.* Potential endpoints for clinical trials in premanifest and early Huntington's disease in the TRACK-HD study: analysis of 24 month observational data. *Lancet Neurol.* **11**, 42–53 (2012).
20. Tabrizi, S. J. *et al.* Predictors of phenotypic progression and disease onset in premanifest and early-stage Huntington's disease in the TRACK-HD study: analysis of 36-month observational data. *Lancet Neurol.* **12**, 637–649 (2013).
21. Klöppel, S. *et al.* Compensation in Preclinical Huntington's Disease: Evidence From the Track-On HD Study. *EBioMedicine* **2**, 1420–1429 (2015).
22. Freeborough, P. A., Fox, N. C. & Kitney, R. I. Interactive algorithms for the segmentation and quantitation of 3-D MRI brain scans. *Comput. Methods Programs Biomed.* **53**, 15–25 (1997).
23. Magnotta, V. A. *et al.* Structural MR image processing using the brains2 toolbox. *Comput. Med. Imaging Graph.* **26**, 251–264 (2002).
24. Schiratti, J.-B., Allassonniere, S., Colliot, O. & Durrleman, S. Learning spatiotemporal trajectories from manifold-valued longitudinal data. Neural Information Processing Systems, Dec 2015, Montréal,

Canada. hal-01163373v3:9.

25. Talairach, J. & Szikla, G. Application of Stereotactic Concepts to the Surgery of Epilepsy. in *Advances in Stereotactic and Functional Neurosurgery 4* (eds. Gillingham, F. J., Gybels, J., Hitchcock, E., Rossi, G. F. & Szikla, G.) 35–54 (Springer, 1980). doi:10.1007/978-3-7091-8592-6_5.
26. Hersch, S. M. *et al.* The CREST-E study of creatine for Huntington disease: A randomized controlled trial. *Neurology* **89**, 594–601 (2017).
27. Reilmann, R. *et al.* Safety and efficacy of pridopidine in patients with Huntington's disease (PRIDE-HD): a phase 2, randomised, placebo-controlled, multicentre, dose-ranging study. *Lancet Neurol.* **18**, 165–176 (2019).
28. McGarry, A. *et al.* A randomized, double-blind, placebo-controlled trial of coenzyme Q10 in Huntington disease. *Neurology* **88**, 152–159 (2017).

Figures



a) Subjects' data

b) Estimation of idealized logistic progression and subjects' parameters

c) Procedure for mapping the data of one subject to the idealized scenario of progression

d) Model fitted to baseline data of a new subject: the predicted progression is compared to follow-up data

Figure 1

HD COURSE MAP describes the changes of clinical and imaging biomarkers in early HD; and then predicts the future values of these biomarkers at the entry of a new participant. a) Schematic representation of two imagined biomarkers (orange and blue) for three subjects of different ages and disease stages (Subject A, B and C). Biomarker values are normalized between 0 and +1, the latter corresponding to the maximum pathological value. Example: subject C shows biomarker values similar to controls at age 60, but the blue biomarker increased to 0.4 at age 72 b) The method estimates two idealized curves for the two biomarkers. Each derives from stitching together the transformed data from these three subjects and using iterative fitting to find best-fit parameters for each biomarker and each subject. The Huntington Age (x axis; HA) of each subject derives from two of these parameters, as described in the text. c) Three parameters describe the relationship of the idealized curves to the actual data: α (acceleration factor); τ (time shift), and ω (intermarker spacing). The modeling assumes that each

subject has a single α, a single τ, and an ω for each biomarker. d) The standard curves are fitted to the baseline data a new subject. The best-fit parameters predict a personalized trajectory of changes of the subject's biomarkers. The predicted values are compared to the subject's data at follow-up visits 1 to 4 years after baseline.

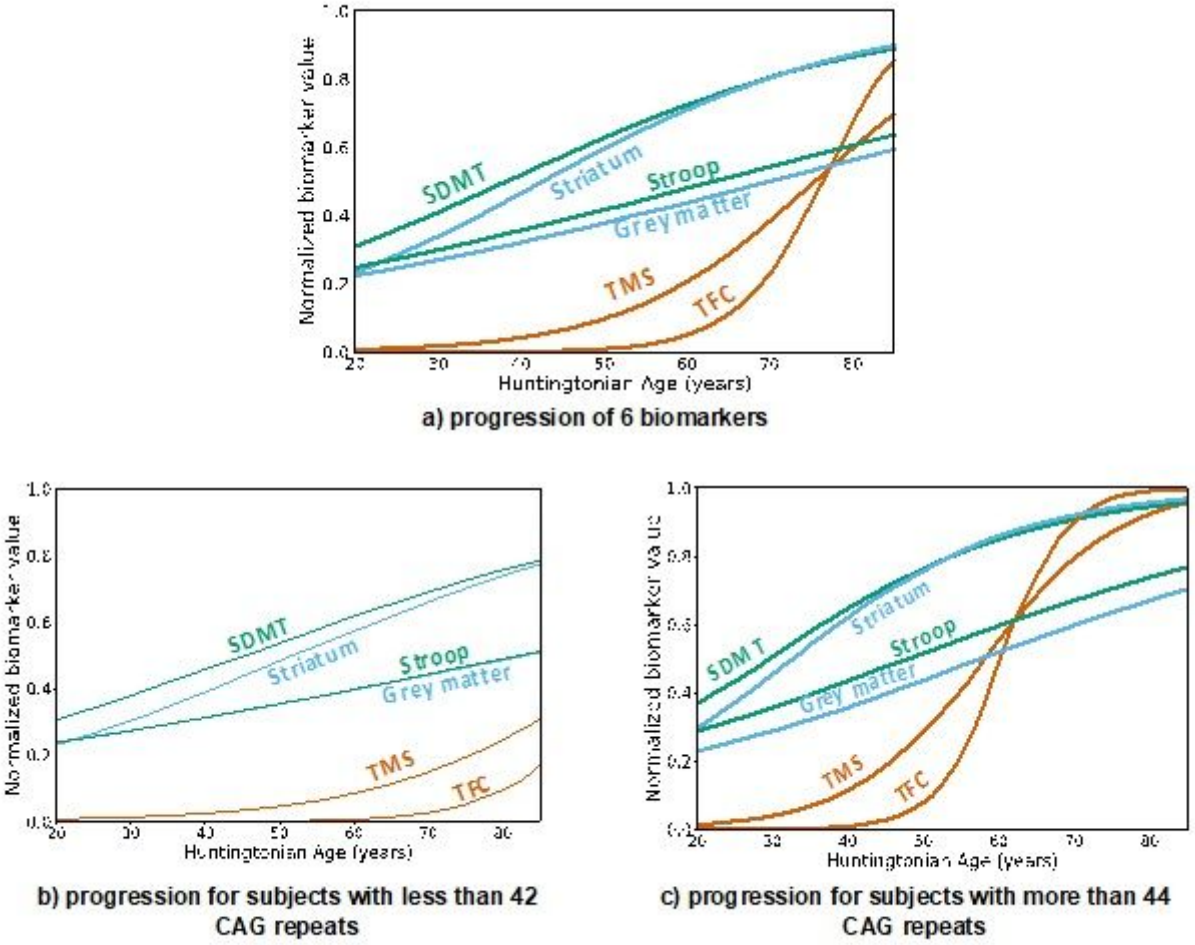


Figure 2

Progression of six out of 15 biomarkers in HD according to the number of CAG repeats. a) The progression estimated from all 299 participants in the TRACK-HD and TRACK-ON studies. Shaded areas correspond to the interquartile confidence intervals around the mean (central curves). Some biomarkers progress in a nearly linear fashion while others exhibit a sharp increase over time. TMS: UHDRS total motor score, Stroop: Stroop Word Reading Test, SDMT: Symbol Digit Modality Test, TFC: UHDRS Total Functional Capacity, Striatum and grey matter: striatal and grey matter volume. b and c) Typical progression for participants with fewer than 42 CAG repeats (b) or more than 44 repeats (c).

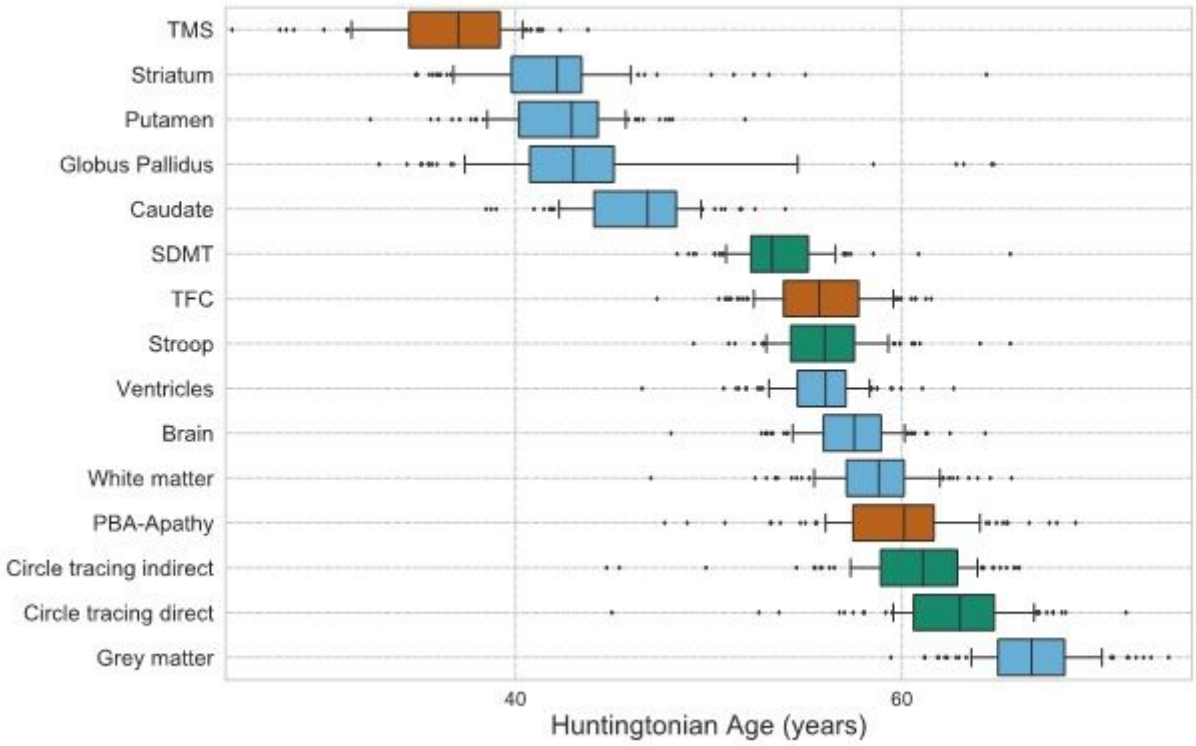


Figure 3

Huntington disease shows a specific temporal ordering of biomarker changes. We repeatedly recalculated the age at which each biomarker diverged from control values (Figure S3). The central line is the median age of 100 simulations. The boxplots show the first and third quartiles, and the whiskers show the first and ninth decile. For instance, motor changes (TMS) are the first detectable sign of HD at HA of 37 years, with an interquartile confidence interval from 34.5 to 39.2 years.

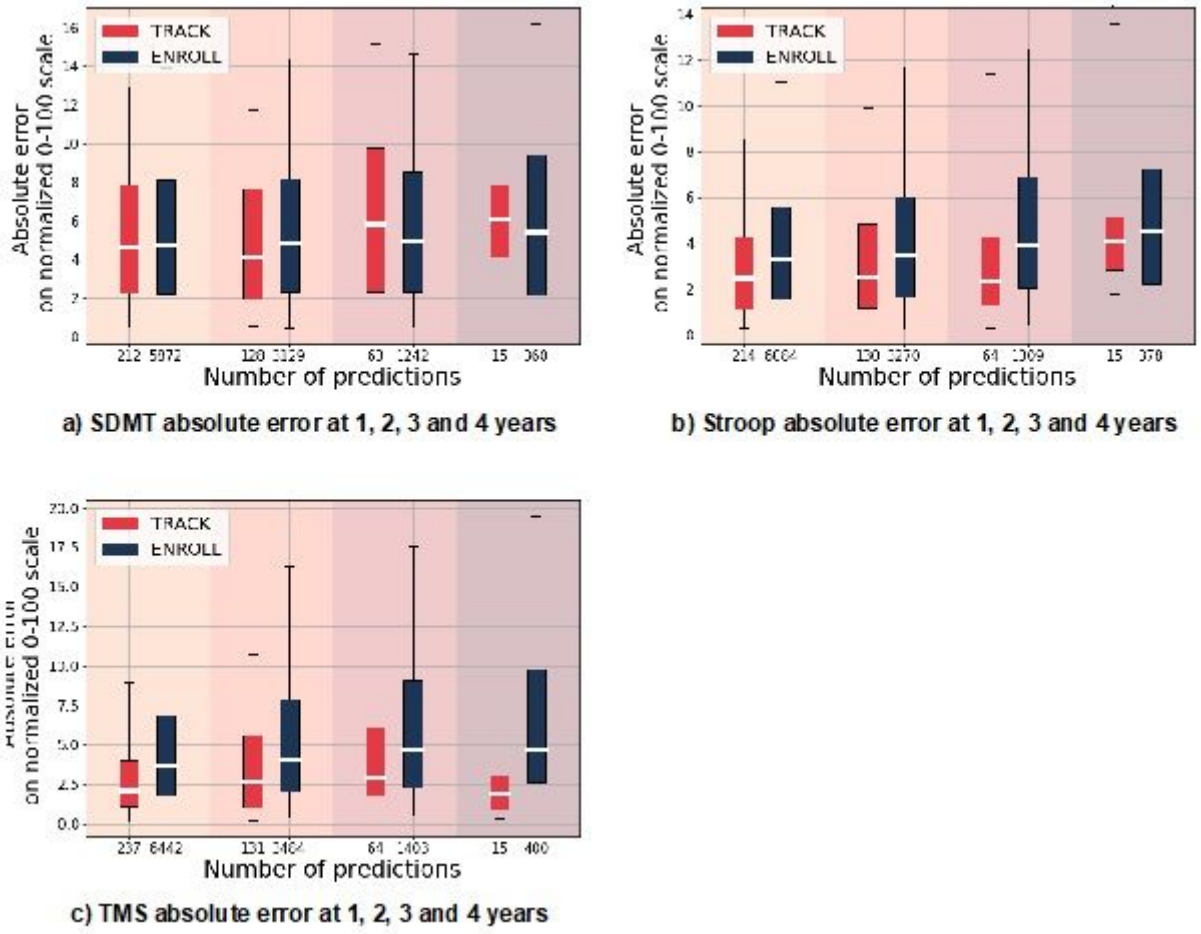


Figure 4

HD COURSE MAP predicts progression at four years in participants from TRACK and ENROLL studies. Prediction of three clinical measurements is shown, SDMT (a), Stroop word reading (b) and TMS (c). Distribution of absolute errors are shown. Prediction on TRACK participants is pictured in red, prediction on ENROLL participants in black. Central squares represent the median and the 25th and 75th percentiles. Whiskers: represent 5th and 95th percentiles. Outliers are removed for clarity. Although HD COURSE MAP was built using TRACK data only, it predicts progression of ENROLL patients with a mean absolute error that does not exceed 5%.

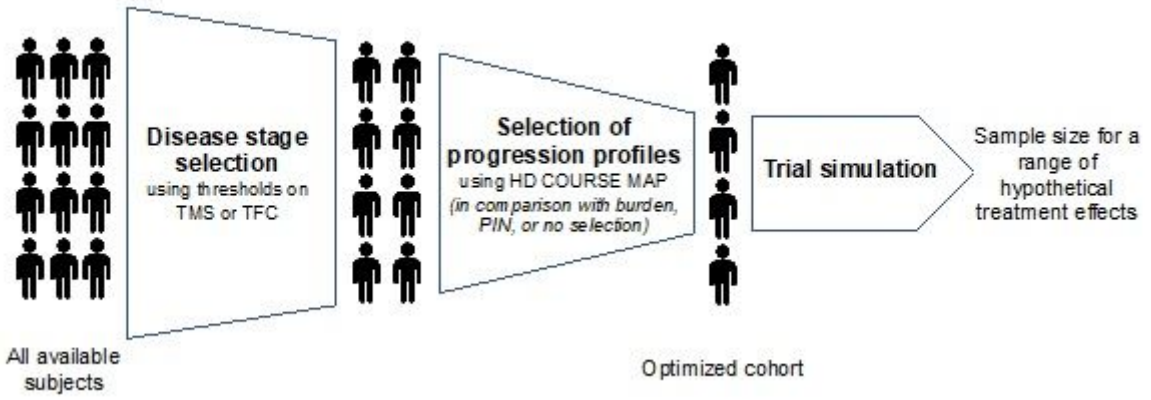


Figure 5

HD COURSE MAP allows the identification of potentially good responders. After selecting participants based on their disease stage, we can further identify participants who will show a significant change during the trial. We have compared the power of hypothetic trials using this approach to hypothetic trials with unselected participants or with selection based on disease burden or PIN scores.

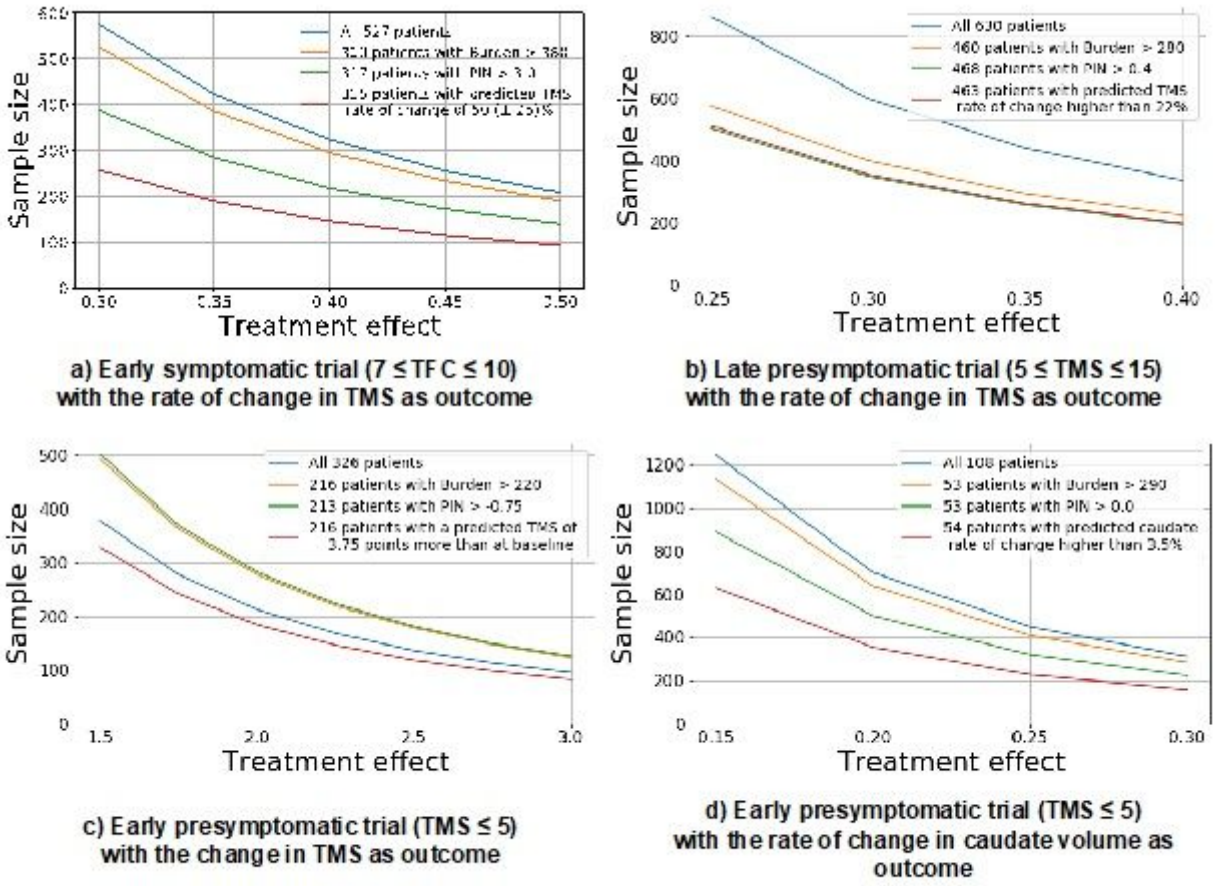


Figure 6

HD COURSE MAP reduces the number of necessary participants regardless of disease stages compared to existing methods. Each panel represents a simulated trial on a particular disease stage with different outcome measures. Blue curves represent every participant meeting the prespecified inclusion criteria. Orange, green and red curves represent participants selected based on the burden score, the PIN score or HD COURSE MAP respectively. The experiments from a to c use ENROLL participants. Experiment in d use TRACK participants.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementarymaterial.docx