RESEARCH ARTICLE OPEN ACCESS

# Modeling Disease Progression and Placebo Response in Huntington Disease

## Insights From Enroll-HD and GENERATION HD1 Cohorts

Marcelo Boareto,[1] Yumi Yamamoto,[1] Jeffrey D. Long,[2,3] Cristina Sampaio,[4,5] Peter McColgan,[6] Cheikh Diack,[1,*] and Patricia Sanwald Ducray[1,*]

*Neurology*® 2025;104:e213646. doi:10.1212/WNL.0000000000213646

**Correspondence**
Dr. Boareto
marcelo.boareto@roche.com

## Abstract

### Background and Objectives

Huntington disease is a rare neurodegenerative disorder with no disease-modifying therapies. This study aimed to quantify longitudinal changes in Unified Huntington's Disease Rating Scale (UHDRS) scores and evaluate their susceptibility to placebo response, enhancing our understanding of disease progression and ability to optimize future trials.

### Methods

We used data from the Enroll-HD natural history study (NCT01574053) and the GENERATION HD1 phase 3 clinical trial (NCT03761849) to model disease progression and placebo response for UHDRS scores, which are commonly used to evaluate disease progression in clinical trials.

### Results

We included 8,071 Enroll-HD participants (mean baseline age: 51.4 years, 51.5% female) to develop a natural history progression model using baseline characteristics as predictive covariates. This model was then used to predict natural history progression of 260 participants from the GENERATION HD1 placebo arm (mean baseline age: 48.7 years, 43.5% female).

The progression measured by Total Functional Capacity (TFC) in the GENERATION HD1 placebo arm aligned with predicted natural history (within 95% CI), indicating no significant placebo response. However, significant improvements (outside the 95% CI of the model) were observed after baseline for Total Motor Score (TMS), Symbol Digit Modalities Test (SDMT) score, and Stroop Word Reading (SWR) score. The improvement in TMS persisted until the end of the dosing period (week 69), converging to natural history predictions at subsequent follow-up visits (weeks 85 and 101), suggesting a placebo effect. By contrast, cognitive scores (SDMT and SWR) showed sustained significant improvement (outside the 95% CI) up to the final follow-up visit at week 101, likely due to practice effects from the more frequent testing schedule in GENERATION HD1 compared with the annual assessments in Enroll-HD. Consequently, the composite UHDRS (cUHDRS) score, a linear combination of TFC, TMS, SDMT, and SWR, is influenced by both placebo and practice effects.

### Discussion

Our results suggest that clinical scores in Huntington disease trials are susceptible to long-term placebo responses. These effects should be considered in future trial designs, especially when comparing trial data with natural history studies. Although based on the largest Huntington's trial, our results rely on limited placebo data and may not generalize to other trials with different populations, treatments, and designs.

MORE ONLINE
**Supplementary Material**



*These authors contributed equally to this work.

[1]Roche Pharma Research and Early Development (pRED), F. Hoffmann-La Roche Ltd., Basel, Switzerland; [2]Department of Psychiatry, University of Iowa, Iowa City; [3]Department of Biostatistics, University of Iowa, Iowa City; [4]CHDI Management, Inc. the company that manages the scientific activities of CHDI Foundation, Princeton, NJ; [5]Laboratorio de Farmacologia Clinica, Faculdade de Medicina de Lisboa (FMUL), Lisbon, Portugal; and [6]Roche Pharma Research and Early Development (pRED), Product Development Neurology, Roche Products, Welwyn Garden City, United Kingdom.

The Article Processing Charge was funded by the authors.

## Glossary

**BMI** = body mass index; **CAG** = cytosine-adenine-guanine; **CAP** = CAG-age product; **CAP** = CAG-age-product; **cUHDRS** = composite UHDRS; **DCL** = diagnostic confidence level; **HD** = Huntington disease; **HD-ISS** = Huntington's Disease Integrated Staging System; **IS** = independence scale; **PIN** = normalized prognostic index; **SDMT** = Symbol Digit Modalities Test; **SWR** = Stroop Word Reading; **TFC** = Total Functional Capacity; **TMS** = Total Motor Score; **UHDRS** = Unified Huntington's Disease Rating Scale.

## Introduction

Disease progression in Huntington disease (HD) is complex and multifaceted, characterized by psychiatric disturbance, motor dysfunction, and cognitive and functional impairment. These changes can be assessed using various clinical scores that are part of the Unified Huntington's Disease Rating Scale (UHDRS),[1] such as the Total Functional Capacity (TFC), Total Motor Score (TMS), Symbol Digit Modalities Test (SDMT), and Stroop Word Reading (SWR). Moreover, these scores can be integrated into a composite UHDRS (cUHDRS) score.[2]

Currently, there is no approved disease-modifying therapy for HD. As with other rare diseases, numerous challenges exist in developing novel treatments, and there is a pressing need to optimize clinical trials in HD. This can be accomplished by more precisely identifying patients who are likely to exhibit disease progression during the trial period and determining the most appropriate end point for a specific patient population. Furthermore, sample size calculations typically rely on real-world data, and the impact of placebo response on the ability to differentiate a treatment effect from a placebo effect is not well understood. Therefore, a deeper understanding and quantification of longitudinal changes in clinical scores and placebo response are important steps in improving the design and analysis of clinical trials for HD.

Mathematical and computational methods have played an important role in advancing our understanding of disease progression in HD. One primary focus of these approaches has been predicting motor onset. The expansion of the cytosine-adenine-guanine (CAG) repeat lengthens the encoded polyglutamine segment of the huntingtin protein. It has been shown that the length of the CAG repeat is a strong predictor of age at motor onset,[3] with various groups characterizing a mathematical relationship between CAG repeat length and the age at motor onset that accounts for over half of its variance.[4-6] Numerous studies have extended the prediction of motor onset beyond CAG repeat length by incorporating longitudinal clinical and imaging data[7] and also by deriving a prognostic measure for predicting disease progression and risk of motor diagnosis in premanifest individuals.[8,9] These studies have significantly improved our prognostic capability concerning age at onset in HD, a relevant achievement, because identifying patients who are likely nearing disease onset is essential for implementing clinical interventions early in the course of the disease, when preventing neuronal death and preserving function are most likely to happen.

In the past years, there has been a growing emphasis on expanding the assessment of HD beyond motor measures. Numerous studies have incorporated biological, clinical, and functional evaluations, including the proposal of a new staging system: the Huntington's Disease Integrated Staging System (HD-ISS).[10] Disease progression models taking into account various aspects of the disease have been proposed. Disease progression in HD has been suggested to happen as a series of states with increasing severity.[11,12] Predictive features of disease progression have been identified,[13-15] and different disease trajectories have been categorized.[15] Moreover, many studies have established the link between changes in features extracted from brain imaging or molecular biomarkers and clinical progression.[16-20] These efforts, along with the availability of large natural history studies such as Enroll-HD,[21] have substantially improved our understanding of disease progression in HD and the key characteristics underlying individual variability in disease progression.

In this study, we developed a mathematical framework to investigate and quantify the natural history progression and placebo response in HD. We use data from Enroll-HD to infer optimal disease trajectories, quantify patient variability in progression rate, and predict progression based on patient characteristics. To gain a deeper understanding of placebo response in HD, we compared the progression in Enroll-HD (annual visits) with data from short-interval follow-up placebo data (GENERATION HD1).[22] Through this comparison, we were able to characterize and quantify placebo response, that is, the variations in disease progression between natural history and the placebo arm of a pivotal trial. Finally, we show that by quantifying both natural history progression and placebo response, we can simulate placebo arm outcomes for patient populations with different inclusion criteria.

## Methods

### Standard Protocol Approvals, Registrations, and Patient Consents

The studies were conducted in accordance with the Declaration of Helsinki. Detailed descriptions of the study protocols have been previously published for the Enroll-HD study (ClinicalTrials.gov Identifier: NCT01574053)[21] and the GENERATION HD1 study (ClinicalTrials.gov Identifier: NCT03761849).[23]

## Study Cohorts

Natural history data used in this work were generously provided by the participants in the Enroll-HD study and made available by CHDI Foundation, Inc. Enroll-HD is a global clinical research platform designed to facilitate clinical research in Huntington disease. We used data from Enroll-HD (NCT01574053) version PDS 6. Our selection criteria included individuals with an HD-ISS Stage greater than 1 and at least one follow-up visit, which was conducted annually. In addition, we selected only those individuals who had all the measurements of the baseline clinical scores and covariates used in the model (in the Covariate Selection section). Less than 3% of potential participants were excluded because of missing covariate measurements, and no imputation methods were applied. This resulted in a total of 8,071 individuals being selected. While participants were followed up to 10 visits, it is important to note that most individuals had fewer than 3 follow-up visits.

To evaluate potential placebo response in HD, we used data from GENERATION HD1 (NCT03761849).[22] This double-blind Phase III study was designed to assess the efficacy and safety of tominersen in patients with HD. The study enrolled 791 adults, with the following overall inclusion criteria: age between 25 and 65 years, diagnostic confidence level (DCL) = 4, independence scale (IS) score ≥70, CAG length ≥36, CAG-age product (CAP) > 400, body mass index (BMI) between 16 and 32 kg/m$^2$, and weight exceeding 40 kg. More details regarding the inclusion and exclusion criteria are available in the protocol (sections 4.1.1 and 4.1.2).[22] Participants were evenly distributed into 3 groups: a placebo group and 2 active groups that received tominersen at 2 different dosing regimens. Patient monitoring occurred every 8 weeks. Based on an overall benefit-risk assessment by an independent data monitoring committee, dosing was halted after most of the patients had completed 69 weeks of the trial. Patients continued to be followed until the 101th week. For the purpose of this analysis, only participants from the placebo group (n = 260) were included. Minimal dropout was observed in the early assessments (approximately 3% at weeks 5 and 21), with dropout rates increasing to approximately 10% at week 69 when dosing was halted and 23% by the final follow-up visit at week 101.

A comparative analysis of the baseline characteristic distribution between both cohorts can be found in eFigure 1 and the study by McColgan et al.[22] Overall, there was good consistency in the distribution of key characteristics and baseline scores between GENERATION HD1 and Enroll-HD participants (eFigure 1). More specifically, GENERATION HD1 participants had a higher education level (EDU) and were less advanced in the disease (as indicated by shorter disease duration and higher TFC scores) compared with Enroll-HD participants. We evaluated the impact of applying the GENERATION HD1 selection criteria to the Enroll-HD cohort. We observed that this adjustment did not significantly alter the overall results. Therefore, no covariate adjustment or statistical balancing between both cohorts was performed in our analysis.

## Mathematical Model

Our approach consists of 3 steps. We first estimate the disease trajectories for each individual score (TFC, TMS, SDMT, and SWR) throughout the course of the disease. For that, we fitted a generalized logistic model[24] with the lower and upper bounds as free parameters, allowing for a wide range of possible trajectories. This approach enabled us to consider multiple trajectory shapes, including exponential, logistic, linear, and combinations thereof, selecting the one that best fit the data. Second, we determine the variability in the rate of progression and identify which covariates are predictive of this progression rate. In both analyses, we use data from Enroll-HD. Finally, we compare natural history progression with GENERATION-HD1 placebo arm data to quantify placebo response.

Our analysis involved specifying a hierarchical Bayesian model. This model incorporated fixed effects for the model structure and covariate coefficients, as well as random effects to account for subject-level variability. To implement our model, we used the probabilistic programming language Turing.jl,[25] which enables us to perform full Bayesian inference. Specifically, we used Hamiltonian Monte Carlo, a Markov chain Monte Carlo method, to generate samples from the posterior distribution. Model parameters can be found in eTables 1–4, and more details are available in eAppendix 1. Hierarchical Bayesian modeling offers a robust approach to handling missing data by integrating uncertainty into the posterior distributions. However, it relies on the assumption that dropout occurs completely at random. While all participants are taken into account when fitting the model parameters, some participants have missing visits. If missing visits are related to the disease status of the participants, this assumption may introduce bias into the analysis.

## Covariate Selection

Patient characteristics at baseline were evaluated as potential predictive covariates for both the progression rate and magnitude of placebo response. The following covariates were included: baseline levels of clinical scores or disease stage (TFC, TMS, SDMT, SWR, IS, HD-ISS), patient demographics (sex, height, weight, BMI, education level, age), co-medication (such as use of tetrabenazine, antidepressant, or antipsychotic), and disease-specific information (CAG, CAP). A linear relationship between the covariates was tested as predictive factors. Nonlinear relationships between the covariates were not addressed in this analysis.

Covariate selection was based on the posterior distribution of each covariate coefficient $\beta_j$. Covariates predictive of progression rate were considered important if the 95% posterior distribution of their coefficient $\beta_j$ did not include 0. For

covariates predictive of placebo response, a threshold of 80% was used. Conversely, if the distribution included 0, the covariates were deemed unimportant and subsequently removed from further analysis. Covariate selection was performed independently for each clinical score.

To address the challenge of testing multiple covariates and account for the inclusion of irrelevant predictors, we used the horseshoe prior.[26,27] This prior was chosen for its ability to mitigate overfitting when assessing a large number of covariates.

A weighted linear combination of the selected covariates was used as a predictor of the progression rate and magnitude of placebo response for each clinical score. More detailed information can be found in eAppendix 1.

## Software

Figure generation and statistical analysis were performed in Julia (version 1.10.5).[28]

## Source Code

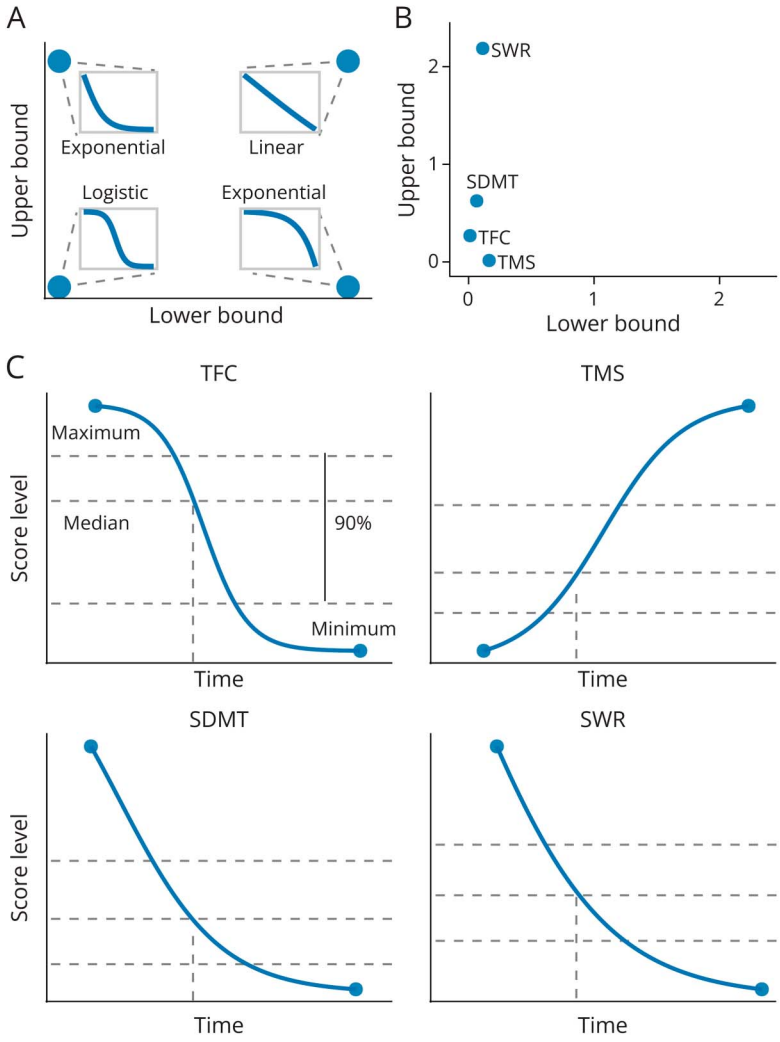
The source code to run the analysis, based on simulated data, can be found at github.com/mboareto/disease_progression_modelling.

## Data Availability

Enroll-HD periodic data set 6 (PDS6) is accessible through the Enroll-HD website (enroll-hd.org/).

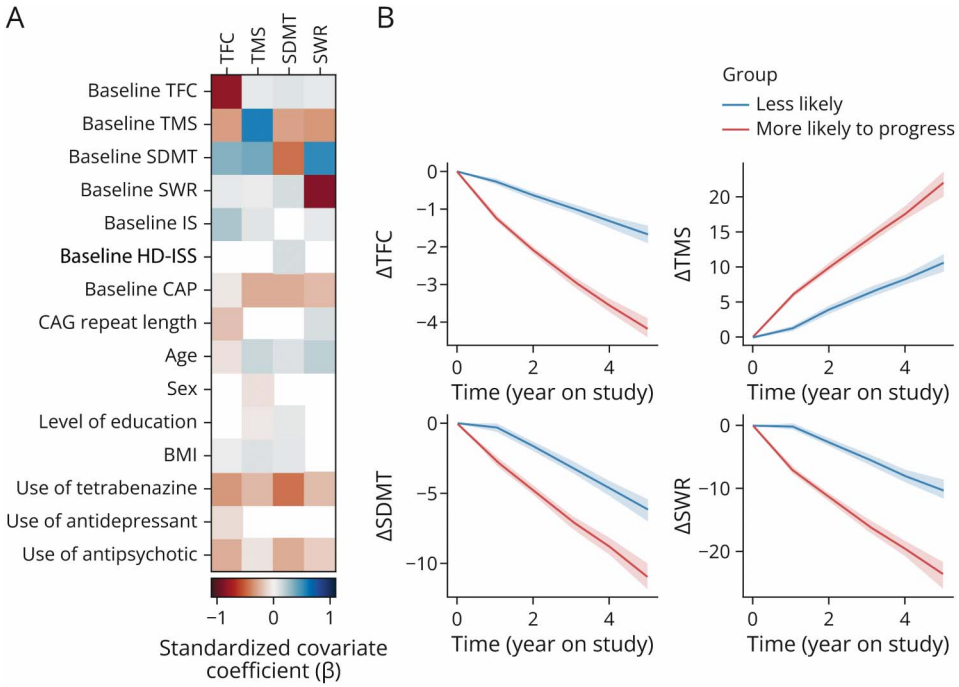
For clinical trial studies, qualified researchers may request access to individual patient-level clinical data through a data request platform. At the time of writing, this request platform is Vivli (vivli.org/ourmember/roche/). Up-to-date details on Roche's Global Policy on the Sharing of Clinical Information and how to request access to related clinical study documents can be found at go.roche.com/data_sharing. Anonymized records for individual patients across more than one data source external to Roche cannot, and should not, be linked because of a potential increase in risk of patient re-identification.

**Figure 1** Differences in Functional, Motor, and Cognitive Trajectories




(A) Schematic illustration of the disease trajectory fitting procedure. By treating the lower and upper bounds of the generalized logistic model as model parameters, we evaluated multiple potential disease trajectories. Inset figures depict extreme examples: exponential, logistic, and linear models. More details can be found in eAppendix 1. (B) TFC and TMS are best characterized by logistic-like trajectories (low lower and low upper bound values) while cognitive scores (SDMT and SWR) are best characterized by exponential-like trajectories (low lower and high upper bound values). (C) Fitted trajectories for TFC, TMS, SDMT score, and SWR score. Dots at the extreme represent minimum and maximum score levels in the data. Dashed lines represent the median and 90% interval. In contrast to other scores, TMS values increase with the progression of the disease. These trajectories provide an average representation of the expected progression over the course of the disease (time). SDMT = Symbol Digit Modalities Test; SWR = Stroop Word Reading; TFC = Total Functional Capacity; TMS = Total Motor Score.

**Figure 2** Baseline Characteristics Can Identify Patients More Likely to Progress



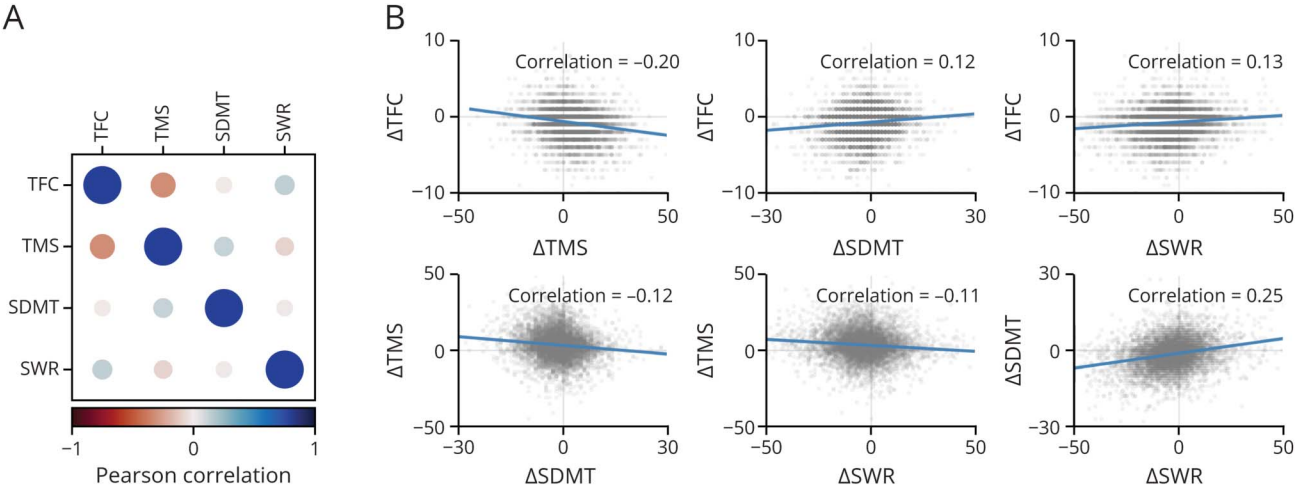

(A) Heatmap illustrates the relevance of each covariate in predicting progression rate for the respective clinical scores. Darker shades indicate a stronger impact. Estimated coefficient values have been standardized according to the average progression rate to enable comparison between different clinical scores. Red shades indicate an increase in the rate of progression while blue shades indicate a decrease. (B) Cross-validation analysis showing changes in TFC, TMS, SDMT score, and SWR score from baseline for individuals selected from Enroll-HD (validation set). Patients were categorized into 50% more likely and 50% less likely to progress based on the model's predictions using baseline covariates. Lines represent the average changes from baseline while the shaded area denotes a 95% CI across 100 cross-validation iterations, using 70% as training data and 30% as validation data from Enroll-HD. SDMT = Symbol Digit Modalities Test; SWR = Stroop Word Reading; TFC = Total Functional Capacity; TMS = Total Motor Score.

# Results

## Disease Progression Trajectories of Functional, Motor, and Cognitive Decline

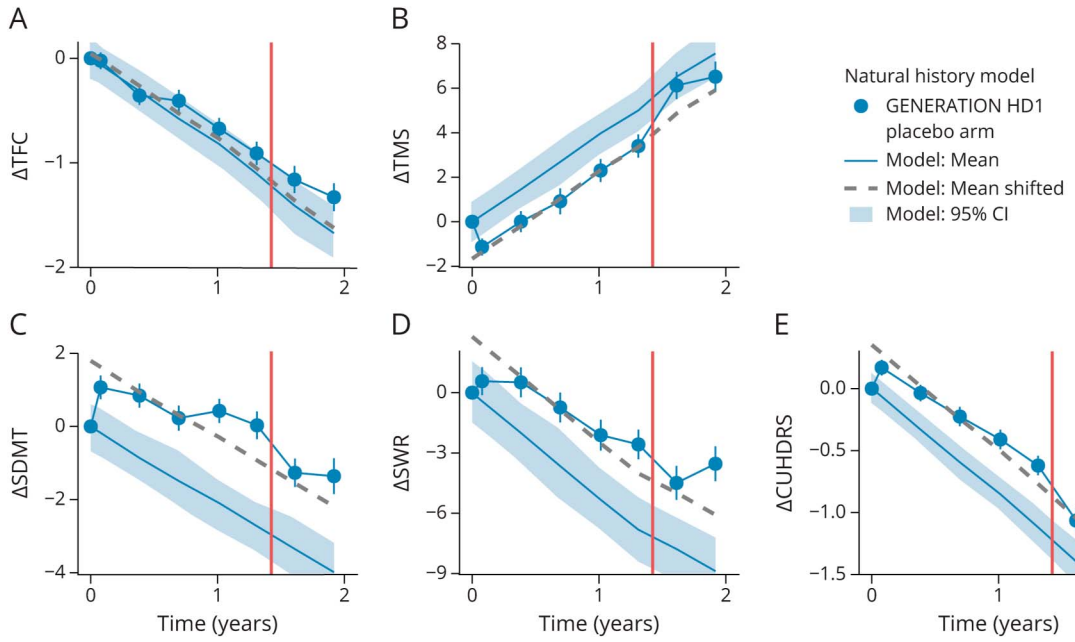
To determine the disease progression trajectories for each clinical score over the course of the disease, we analyzed natural history data from Enroll-HD. We screened for multiple trajectory shapes, including exponential, logistic, linear, and combinations thereof, selecting the one that best fit the data (more details in eAppendix 1 and Figure 1A). This method was independently applied to each UHDRS clinical score (TFC, TMS, SDMT, SWR). Our analysis revealed that TFC and TMS are best characterized by logistic-like trajectories while cognitive scores (SDMT and SWR) display an

**Figure 3** Correlation Among Rates of Functional, Motor, and Cognitive Decline




(A) Pearson correlation coefficients for the predicted progression rates among TFC, TMS, SDMT score, and SWR score. (B) Associations between the one-year changes in TFC, TMS, SDMT score, and SWR score for individuals selected from the Enroll-HD database. Gray dots represent the change in each clinical score after 1 year for individual patients while the blue line illustrates the linear fit of the relationship between changes in each score. Insets in the top right corner of each subplot display the correlation coefficients for the respective score changes. SDMT = Symbol Digit Modalities Test; SWR = Stroop Word Reading; TFC = Total Functional Capacity; TMS = Total Motor Score.

**Figure 4** Placebo Response on Functional, Motor, and Cognitive Scores




This figure illustrates the changes from baseline in (A) TFC, (B) TMS, (C) SDMT score, (D) SWR score, and (E) cUHDRS score. Blue dots and lines indicate the average data from the GENERATION HD1 placebo arm while vertical blue lines represent the standard error. The vertical red line marks the end of the dosing period for most patients (after week 69). The blue line shows the prediction based on the natural history model, with the shaded area representing the 95% CI. The dashed line indicates a shift from the mean predictions of the model, demonstrating that by adjusting the model prediction proportionally to an initial improvement, the progression rate in GENERATION HD1 aligns with the predicted natural history. cUHDRS = composite UHDRS; SDMT = Symbol Digit Modalities Test; SWR = Stroop Word Reading; TFC = Total Functional Capacity; TMS = Total Motor Score.

exponential-like decline (Figure 1, B and C). These findings indicate that different clinical scores evolve at distinct rates over the course of the disease.

## Predicting Rate of Clinical Decline

After defining disease trajectories for each clinical measurement, we further explored whether baseline patient characteristics could predict the rate at which participants from the Enroll-HD cohort progress along these predefined trajectories. Previous studies have shown that baseline score levels, CAP, CAG length, and co-medication are predictive of progression rate.[13,14] Based on these studies, we selected 15 patient characteristics representing disease stage, patient demographics, co-medication use, and disease-specific information as potential predictive covariates to be evaluated (Figure 2A, Methods).

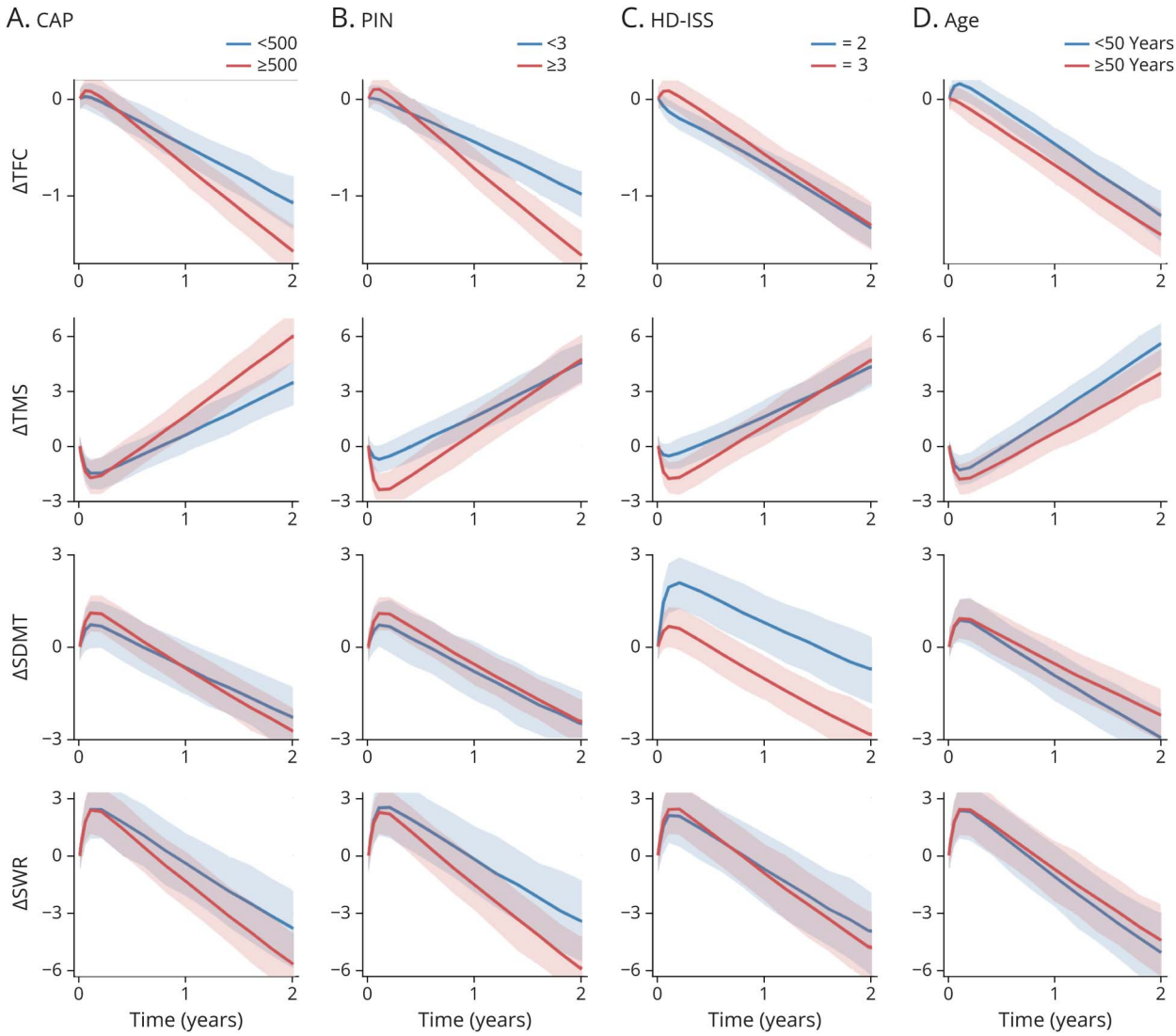
Many covariates exhibit high correlations with each other because they often represent similar aspects of the disease. For instance, baseline score levels indicate the stage of the disease and are, therefore, highly correlated. This poses limitations when interpreting the relevance of a covariate based on their coefficients. It is important to note that a low coefficient for a correlated covariate does not imply its irrelevance because its effect might be represented by other covariates. Keeping these limitations in mind, we observed that, overall, the most relevant covariates include the stage of the disease (HD-ISS and baseline clinical scores), disease-specific characteristics (CAG, CAP), and the use of tetrabenazine and antipsychotics while patient demographics (sex, BMI, education, age) have small relevance.

To validate the predictive nature of these baseline characteristics on progression rates, we performed a cross-validation analysis by randomly dividing the Enroll-HD data set into a 70%–30% training-validation set and repeating this process 100 times. We used the training set to estimate the relevance of each covariate and then categorized patients in the validation set into 2 groups, the 50% more likely to progress and 50% less likely to progress based on their predicted progression rates. We observed a strong separation between these 2 predicted populations, indicating that baseline characteristics serve as strong predictors of progression rate and can effectively identify patients who are more likely to experience disease progression (Figure 2B).

## Correlation in Progression Rates Among Different Clinical Scores

In our initial analysis, we found that functional, motor, and cognitive scores are best characterized by distinct disease trajectories (Figure 1), implying that patients may exhibit varying rates of functional, motor, and cognitive decline throughout the course of the disease. To further examine this, we estimated the correlation of predicted progression rates at baseline (time of entry of the study) among different clinical scores. Our model predicts a weak correlation in the progression

**Figure 5** Simulation of Placebo Arms for Different Patient Populations




Changes from baseline in TFC, TMS, SDMT score, and SWR score for patients segregated by (A) HD-ISS, (B) PIN, (C) CAP, and (D) age. For each scenario, we randomly sampled 250 participants from Enroll-HD for each condition and used their baseline characteristics to simulate the predicted changes in the clinical score. We performed 1,000 random samplings. The solid line represents the average while the shaded area depicts the 95% CI prediction. For the segregation of the plots into 2 groups, we chose values of CAP, PIN, and age that are close to the median of the Enroll-HD population (493 for CAP, 3.07 for PIN, and 51 for age). CAP = CAG-age-product; PIN = normalized prognostic index; SDMT = Symbol Digit Modalities Test; SWR = Stroop Word Reading; TFC = Total Functional Capacity; TMS = Total Motor Score.

rates between these clinical scores (Figure 3A). By plotting the relationship between short-term changes in clinical scores from individuals selected from the Enroll-HD cohort, we further confirmed that changes in functional, motor, and cognitive scores are weakly correlated (Figure 3B). These results highlight the temporal differences in the rate of progression captured by different clinical end points and emphasize that individuals who exhibit rapid progression according to one clinical score may not necessarily exhibit fast progression according to another.

## Placebo Response in GENERATION HD1

To quantify the placebo response, we compared the progression in the GENERATION HD1 placebo cohort with their expected natural history progression. We observed that for the functional score (TFC), the progression during the trial aligns with the natural history progression, showing no sign of a relevant placebo response (Figure 4A). By contrast, for the motor and cognitive measurements, we found that patients in the placebo arm of GENERATION HD1 have better outcomes than expected by natural history progression (Figure 4, B–D). We also noted that this improvement has a fast onset after treatment initiation and persists throughout the dosing period of the trial. Moreover, after this initial improvement, motor and cognitive decline follows the expected rate of decline from natural history. This finding indicates the presence of a placebo response that contributes to a noticeable improvement in motor and cognitive measurements within a few weeks after the start of the trial, and this improvement persists consistently throughout the trial duration.

Intriguingly, we identified differences in the placebo response between motor and cognitive scores. For the Total Motor Score (TMS), the improvement lasts up to the end of the dosing regimen of the trial. After the dosing regimen ends at week 69 for most participants, the values of TMS align with the expected natural history progression (Figure 4B). However, this was not the case for the cognitive measurements, where we observed the persistence of this initial improvement after dosing was stopped (Figure 4, C and D). The placebo response observed in motor and cognitive scores affect the cUHDRS score because it is derived from a linear combination of TFC, TMS, SDMT score, and SWR score (Figure 4E).

It is important to note that the minimal dropout rate of approximately 3% observed at early assessments (weeks 5 and 21) suggests that the initial improvement after baseline is not influenced by dropout effects. This low dropout rate allows for a robust estimation of the parameters associated with the placebo response, which is modeled as an initial improvement after baseline and, therefore, mostly dependent on the differences between baseline and initial follow-up visits. Consequently, our placebo response model effectively captured the changes in clinical scores within the GENERATION HD1 placebo cohort during the trial's dosing period (up to week 69) (eFigure 2). Although dropout rates increased to approximately 10% at week 69 and 23% by week 101, there were no significant differences in baseline characteristics between participants who dropped out and those who completed the trial at week 101 (eFigure 3). However, dropouts progressed faster according to some end points (eFigure 4), which could influence later observations and potentially explain the slightly slower progression observed in some end points, particularly in the cUHDRS score (Figure 4E).

### Simulation of Placebo Arms for Patient Populations With Different Inclusion Criteria

Once natural history progression, placebo response, and baseline characteristics predictive of progression rate are quantified, it is possible to simulate the changes in each clinical score for different patient populations characterized by different baseline characteristics. In Figure 5, we present the expected changes in each clinical measurement for commonly used selection criteria in HD: CAG-age product (CAP), normalized prognostic index (PIN), age, and HD-ISS (different subpopulations given in eFigures 5–8).

We observed variations in the progression of different subpopulations depending on the specific clinical measurement used. For instance, PIN and CAP can effectively enrich the population with patients more likely to progress, particularly when using TFC (Figure 5, A and B). Moreover, when using TFC as an end point, the progression rate in individuals with HD-ISS = 2 is not expected to be different from the rate in those with HD-ISS = 3, in contrast to other clinical measurements such as SMDT (Figure 5C). Finally, no relevant difference is observed when selecting by age (Figure 5D). These results highlight the importance for careful selection of appropriate clinical end points and consideration of specific subpopulations when designing and analyzing clinical trials for HD.

## Discussion

In this study, we aimed to understand and quantify disease progression and placebo response in the UHDRS clinical measurements used to assess functional (TFC), motor (TMS), and cognitive (SDMT, SWR) progression in HD. We used large natural history cohort data from Enroll-HD to model the natural history progression of these clinical scores. We developed a method to estimate disease trajectory for each clinical score and found that TFC and TMS were best characterized by logistic-like trajectories, consistent with previous analysis,[13] while cognitive scores (SDMT, SWR) exhibited exponential-like trajectories.

We further investigated whether baseline characteristics that can be easily assessed in a clinical visit could serve as predictors of the progression rates of each clinical score. Our findings indicate that baseline characteristics such the stage of the disease (baseline levels of clinical scores), CAG, CAP, and use of tetrabenazine and antipsychotics act as robust predictors of progression rate for each clinical score. It is important to highlight that our framework does not establish causality between these characteristics and disease progression. For instance, we cannot determine whether comedication use causes faster progression or whether those with faster progression are more likely to use these medications. In addition, when comparing the progression rates among different scores over 1 year, we observed a weak correlation among them, indicating that patients who are likely to experience progression in one clinical score may not necessarily exhibit the same progression in another score in the short term.

Furthermore, we sought to understand the nature and magnitude of placebo response by analyzing data from the GENERATION HD1 placebo arm. Our analysis revealed minimal placebo response in the functional score (TFC). By contrast, we observed a strong placebo response in motor and cognitive scores, showing a strong improvement after baseline that persisted until at least week 69 when the dosing regimen of the trial stopped for most participants. For the motor score (TMS), after week 69, this improvement diminished, and TMS levels aligned with the expected levels from natural history. This pattern suggests the presence of a long-term placebo effect in the motor score. Indications of placebo effects in TMS have been previously observed in other HD trials,[29-31] and placebo effects in motor scores have been documented in Parkinson disease.[32-35] Of interest, the cognitive scores (SDMT, SWR) exhibited an even stronger placebo response that was sustained after the dosing period of the trial ended. There are several potential causes for the observed strong improvement in the cognitive scores, such as

placebo effect, practice effect, or improvements in mood due to trial participation. Although we cannot exclude any of these possibilities, the observation that this improvement remains at weeks 85 and 101 (after the end of the dosing regimen) suggests that this may be related to a learning or practice effect. This effect is likely influenced by the more frequent assessments in the GENERATION HD1 trial compared with the annual assessments in the Enroll-HD cohort. This is consistent with previous studies showing that the frequency of assessment of cognitive scores lead to improvements related to practice effect in UHDRS cognitive scores,[36,37] but also in cognitive measurements in Alzheimer and multiple sclerosis.[38-40] It is important to note that these effects also significantly influence the cUHDRS score, given that it is derived from a linear relationship between TFC, TMS, SDMT score, and SWR score.[2]

Taken together, these findings have significant implications for trial design in HD. The strong predictive value of baseline characteristics for progression rate allows for the selection of patients based on their expected rate of progression, enabling, for example, trial enrichment with patients more likely to progress over the period of the trial. However, the weak correlation in progression rates between different clinical scores indicates that patients likely to progress based on one score may not necessarily progress based on other scores. This is because the clinical scores change in a nonlinear manner, that is, with different speeds at different moments of the disease. Therefore, depending on where patients are in the course of the disease, they are more likely to progress according to one end point but not another. This highlights the importance of carefully selecting end points when enriching trials with specific patient populations. Furthermore, the observed differences in the nature and magnitude of the placebo response highlight the need for a better understanding of their underlying effects. Moreover, previous research has shown that a composite measure of functional, motor, and cognitive scores (cUHDRS) provides an improved measure of clinical progression and enhances clinical trial design by requiring smaller sample sizes.[2] The use of this composite score is becoming common practice; however, because these analyses were evaluated in the absence of trial data, future work should evaluate whether these conclusions hold in the presence of placebo response.

It is important to note that our analysis revealed that a placebo response initially leads to an improvement after baseline, but after accounting for this initial improvement, the rate of progression between the GENERATION HD1 placebo arm and Enroll-HD cohorts remained consistent. This highlights the value of Enroll-HD as a retrospective source of natural history information. However, our analysis also emphasizes the potential for misleading interpretations when comparing trial data with retrospective natural history studies, particularly in distinguishing placebo response from drug effect.

Our mathematical framework represents an initial step toward quantifying changes in clinical end points within a clinical trial setting. By incorporating both natural history progression inferred from Enroll-HD and placebo response inferred from GENERATION HD1, our model effectively captures changes in clinical end points for different patient populations and can be used to optimize the design and analysis of clinical trials in HD. Moreover, our model can serve as the first step toward having a modeling tool for simulating virtual placebo arms, which can optimize trials by reducing the number of required patients and accelerating clinical research.

It is important to acknowledge the main limitation of the current model, which lies in its reliance on a limited amount of trial data (260 participants from GENERATION-HD1 placebo arm). Therefore, the observed placebo response in GENERATION-HD1 may not necessarily be generalized to other trials with different patient populations, treatment interventions, and trial designs. Moreover, while Enroll-HD includes participants from various stages of the disease, it is worth noting that certain stages, such as early stages, may be less represented. If future trials focus on these specific populations, it may be necessary to adapt the model accordingly to ensure its applicability.

In summary, our findings bring a new understanding of disease progression and placebo response in HD and our theoretical framework has the potential to enhance the efficiency of clinical trials and accelerate the development of effective treatments for patients with HD.

## Acknowledgment

Natural history data used in this work were generously provided by the participants in the Enroll-HD study and made available by CHDI Foundation, Inc. Enroll-HD is a clinical research platform and longitudinal observational study for families with Huntington disease intended to accelerate progress toward therapeutics; it is sponsored by CHDI Foundation, a nonprofit biomedical research organization exclusively dedicated to collaboratively developing therapeutics for HD. Enroll-HD would not be possible without the vital contribution of the research participants and their families. The authors thank all CHDI/Enroll-HD participants and all the participants in GENERATION HD1, their families, the site staff, and the trial team, for their time and commitment to the trial, and the following colleagues and HD experts for their critical feedback on method development and discussions: Francesco Brizzi, Tony Kam-Thong, Antoine Soubret, Ben Ribba, and David Hawellek.

## Author Contributions

M. Boareto: drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data. Y. Yamamoto: study concept or design; analysis or interpretation of data. J.D. Long: analysis or interpretation of data. C. Sampaio: analysis or interpretation of data. P. McColgan: study concept or design; analysis or interpretation of data. C. Diack: study concept or design; analysis or interpretation of data. P.S. Ducray: study concept or design; analysis or interpretation of data.

## Study Funding

Supported by F. Hoffmann-La Roche.

## Disclosure

M. Boareto and Y. Yamamoto are full-time employees of F. Hoffmann-La Roche. J.D. Long reports grants from NIH and CHDI; consulting fees from Alnylam, Alexion, Annexon, AskBio, LifeEdit, Neurocrine, Prilenia, PTC Therapeutics Inc., F. Hoffmann-La Roche, Spark Therapeutics, Teva, uniQure Biopharma, Vertex, and Wave. C. Sampaio is an employee of CHDI Management Inc. advisors for CHDI Foundation, and has received consultancy honoraria from Pfizer, Kyowa Kirin, vTv Therapeutics, GW Pharmaceuticals, Neuraly, Neuroderm, Neuroxpharm, Inflictis, Biocodex, Thelonious Mind, Novartis, Biogen Green Valley Pharmaceuticals, and Pinteeon Pharmaceuticals. P. McColgan, C. Diack, and P.S. Ducray are full-time employees of F. Hoffmann-La Roche. Go to Neurology.org/N for full disclosures.

## Publication History

Previously published in medRxiv (doi.org/10.1101/2024.02.15.24302445). Received by *Neurology*® October 18, 2024. Accepted in final form March 14, 2025. Submitted and externally peer reviewed. The handling editor was Associate Editor Peter Hedera, MD, PhD.

## References

1. Kieburtz K, Penney JB, Como P, et al. Unified Huntington's disease rating scale: reliability and consistency. *Mov Disord*. 1996;11(2):136-142. doi:10.1002/mds.870110204
2. Schobel SA, Palermo G, Auinger P, et al. Motor, cognitive, and functional declines contribute to a single progressive factor in early HD. *Neurology*. 2017;89(24):2495-2502. doi:10.1212/WNL.0000000000004743
3. Andrew SE, Goldberg YP, Kremer B, et al. The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. *Nat Genet*. 1993;4(4):398-403. doi:10.1038/ng0893-398
4. Langbehn DR, Hayden MR, Paulsen JS, the PREDICT-HD Investigators of the Huntington Study Group. CAG-repeat length and the age of onset in Huntington disease (HD): a review and validation study of statistical approaches. *Am J Med Genet B Neuropsychiatr Genet*. 2010;153B(2):397-408. doi:10.1002/ajmg.b.30992
5. Zhang Y, Long JD, Mills JA, Warner JH, Lu W, Paulsen JS, PREDICT-HD Investigators and Coordinators of the Huntington Study Group. Indexing disease progression at study entry with individuals at-risk for Huntington disease. *Am J Med Genet Part B Neuropsychiatr Genet*. 2011;156B(7):751-763. doi:10.1002/ajmg.b.31232
6. Warner JH, Long JD, Mills JA, et al. Standardizing the CAP score in Huntington's disease by predicting age-at-onset. *J Huntington's Dis*. 2022;11(2):153-171. doi:10.3233/jhd-210475
7. Paulsen JS, Long JD, Ross CA, et al. Prediction of manifest Huntington's disease with clinical and imaging measures: a prospective observational study. *Lancet Neurol*. 2014;13(12):1193-1201. doi:10.1016/S1474-4422(14)70238-8
8. Long JD, Langbehn DR, Tabrizi SJ, et al. Validation of a prognostic index for Huntington's disease. *Mov Disord*. 2017;32(2):256-263. doi:10.1002/mds.26838
9. Langbehn DR, Fine EM, Meier A, Hersch S. Utility of the Huntington's disease prognostic index score for a perimanifest clinical trial. *Mov Disord*. 2022;37(5):1040-1046. doi:10.1002/mds.28944
10. Tabrizi SJ, Schobel S, Gantman EC, et al. A biological classification of Huntington's disease: the integrated staging system. *Lancet Neurol*. 2022;21(7):632-644. doi:10.1016/S1474-4422(22)00120-X
11. Sun Z, Ghosh S, Li Y, et al. A probabilistic disease progression modeling approach and its application to integrated Huntington's disease observational data. *JAMIA Open*. 2019;2(1):123-130. doi:10.1093/jamiaopen/ooy060
12. Mohan A, Sun Z, Ghosh S, et al. A machine-learning derived Huntington's disease progression model: insights for clinical trial design. *Mov Disord*. 2022;37(3):553-562. doi:10.1002/mds.28866
13. Sun W, Zhou D, Warner JH, Langbehn DR, Hochhaus G, Wang Y. Huntington's disease progression: a population modeling approach to characterization using clinical rating scales. *J Clin Pharmacol*. 2020;60(8):1051-1060. doi:10.1002/jcph.1598
14. Ghazaleh N, Houghton R, Palermo G, Schobel SA, Wijeratne PA, Long JD. Ranking the predictive power of clinical and biological features associated with disease progression in Huntington's disease. *Front Neurol*. 2021;12:678484. doi:10.3389/fneur.2021.678484
15. Ko J, Furby H, Ma X, et al. Clustering and prediction of disease progression trajectories in Huntington's disease: an analysis of Enroll-HD data using a machine learning approach. *Front Neurol*. 2022;13:1034269. doi:10.3389/fneur.2022.1034269
16. Tabrizi SJ, Reilmann R, Roos RA, 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*. 2012;11(1):42-53. doi:10.1016/S1474-4422(11)70263-0
17. Wijeratne PA, Garbarino S, Gregory S, et al. Revealing the timeline of structural MRI changes in premanifest to manifest Huntington disease. *Neurol Genet*. 2021;7(5):e617. doi:10.1212/NXG.0000000000000617
18. Abeyasinghe PM, Long JD, Razi A, et al. Tracking Huntington's disease progression using motor, functional, cognitive, and imaging markers. *Mov Disord*. 2021;36(10):2282-2292. doi:10.1002/mds.28650
19. Rodrigues FB, Byrne LM, Tortelli R, et al. Mutant Huntingtin and neurofilament light have distinct longitudinal dynamics in Huntington's disease. *Sci translational Med*. 2020;12(574):eabc2888. doi:10.1126/scitranslmed.abc2888
20. Byrne LM, Rodrigues FB, Blennow K, et al. Neurofilament light protein in blood as a potential biomarker of neurodegeneration in Huntington's disease: a retrospective cohort analysis. *Lancet Neurol*. 2017;16(8):601-609. doi:10.1016/S1474-4422(17)30124-2
21. Landwehrmeyer GB, Fitzer-Attas CJ, Giuliano JD, et al. Data analytics from Enroll-HD, a global clinical research platform for Huntington's disease. *Mov Disord Clin Pract*. 2017;4(2):212-224. doi:10.1002/mdc3.12388
22. McColgan P, Thobhani A, Boak L, et al. Tominersen in adults with manifest Huntington's disease. *N Engl J Med*. 2023;389(23):2203-2205. doi:10.1056/nejmc2300400
23. Tabrizi SJ, Leavitt BR, Landwehrmeyer GB, et al. Targeting huntingtin expression in patients with Huntington's disease. *N Engl J Med*. 2019;380(24):2307-2316. doi:10.1056/nejmoa1900907
24. Stukel TA. Generalized logistic models. *J Am Stat Assoc*. 1988;83(402):426-431. doi:10.1080/01621459.1988.10478613
25. Ge H, Xu K, Ghahramani Z. Turing: a language for flexible probabilistic inference. *International Conference on Artificial Intelligence and Statistics*. PMLR; 2018:1682-1690.
26. Carvalho CM, Polson NG, Scott JG. The horseshoe estimator for sparse signals. *Biometrika*. 2010;97(2):465-480. doi:10.1093/biomet/asq017
27. Zhang YD, Naughton BP, Bondell HD, Reich BJ. Bayesian regression using a prior on the model fit: the R2-D2 shrinkage prior. *J Am Stat Assoc*. 2022;117(538):862-874. doi:10.1080/01621459.2020.1825449
28. Bezanson J, Edelman A, Karpinski S, Shah VB. Julia: a fresh approach to numerical computing. *SIAM Rev*. 2017;59(1):65-98. doi:10.1137/141000671
29. de Yebenes JG, Landwehrmeyer B, Squitieri F, et al. Pridopidine for the treatment of motor function in patients with Huntington's disease (MermaiHD): a phase 3, randomised, double-blind, placebo-controlled trial. *Lancet Neurol*. 2011;10(12):1049-1057. doi:10.1016/S1474-4422(11)70233-2
30. Reilmann R, McGarry A, Grachev ID, 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*. 2019;18(2):165-176. doi:10.1016/S1474-4422(18)30391-0
31. Reilmann R, Anderson KE, Feigin A, et al. Safety and efficacy of laquinimod for Huntington's disease (LEGATO-HD): a multicentre, randomised, double-blind, placebo-controlled, phase 2 study. *Lancet Neurol*. 2024;23(3):243-255. doi:10.1016/S1474-4422(23)00454-4
32. de la Fuente-Fernández R, Stoessl AJ. The placebo effect in Parkinson's disease. *Trends Neurosci*. 2002;25(6):302-306. doi:10.1016/S0166-2236(02)02181-1
33. Goetz CG, Leurgans S, Raman R, Stebbins GT. Objective changes in motor function during placebo treatment in PD. *Neurology*. 2000;54(3):710-714. doi:10.1212/WNL.54.3.710
34. Diederich NJ, Goetz CG. The placebo treatments in neurosciences: new insights from clinical and neuroimaging studies. *Neurology*. 2008;71(9):677-684. doi:10.1212/01.wnl.0000324635.49971.3d
35. Espay AJ, Norris MM, Eliassen JC, et al. Placebo effect of medication cost in Parkinson disease: a randomized double-blind study. *Neurology*. 2015;84(8):794-802. doi:10.1212/WNL.0000000000001282
36. Beglinger LJ, Gaydos B, Tangphao-Daniels O, et al. Practice effects and the use of alternate forms in serial neuropsychological testing. *Arch Clin Neuropsychol*. 2005;20(4):517-529. doi:10.1016/j.acn.2004.12.003
37. Beglinger LJ, Adams WH, Fiedorowicz JG, et al. Practice effects and stability of neuropsychological and UHDRS tests over short retest intervals in Huntington disease. *J Huntington's Dis*. 2015;4(3):251-260. doi:10.3233/jhd-150159
38. Goldberg TE, Harvey PD, Wesnes KA, Snyder PJ, Schneider LS. Practice effects due to serial cognitive assessment: implications for preclinical Alzheimer's disease randomized controlled trials. *Alzheimers Dement (Amst)*. 2015;1(1):103-111. doi:10.1016/j.dadm.2014.11.003
39. Roar M, Illes Z, Sejbaek T. Practice effect in symbol digit modalities test in multiple sclerosis patients treated with natalizumab. *Mult Scler Relat Disord*. 2016;10:116-122. doi:10.1016/j.msard.2016.09.009
40. Fuchs TA, Gillies J, Jaworski MG 3rd, et al. Repeated forms, testing intervals, and SDMT performance in a large multiple sclerosis dataset. *Mult Scler Relat Disord*. 2022;68:104375. doi:10.1016/j.msard.2022.104375