Sample Enrichment Strategies for Prodromal Huntington's Disease Trials: Power and Generalizability

May 3, 2024

Ben Bodek

Under the advisement of Dr. Tanya Garcia

Abstract

Huntington's Disease (HD) is a rare genetic neurodegenerative disease. For patients with the genetic mutation that causes HD, symptoms can progress very slowly over the course of years or decades. Thus it is difficult to design clinical trials which are adequately powered to detect a clinically significant slowing of disease progression. "Sample Enrichment strategies" aim to increase clinical trial power by recruiting participants who are expected to experience faster disease progression. This study used a flexible simulation based sample size estimation procedure to compare the effect of sample enrichment strategies on the sample size for an HD clinical triad required to reach 90% power to detect a 30% treatment effect. The evaluated sample enrichment strategies were shown to lead to much smaller sample size requirements. The generalizability of sample enrichment based clinical trials for HD was also investigated by comparing demographic and clinical characteristics of participants in the ENROLL-HD observational study who would or would not be eligible for a clinical trial under each enrichment strategy.

1 Introduction

Huntington's Disease (HD) is a rare inheritable neurodegenerative disease, caused by an excessive number of CAG trinucleotide repeats on the HTT gene (>36 CAG repeats). HD is an autosomal dominant disease, meaning that any individual possessing the expanded CAG mutation on the HTT gene will eventually develop HD[1]. Motor abnormalities are the primary diagnostic criteria for HD. Patients undergo 'motor onset' and are diagnosed with manifest HD when an expert determines that there is a $\geqslant 99\%$ likelihood that their motor abnormalities are caused by HD[2]. Patients with expanded CAG repeats on the HTT gene who have not yet reached motor diagnosis are said to be in the 'Premanifest' stage of HD[2]. Those experiencing measurable motor and cognitive symptoms but who have not yet undergone motor diagnosis are further classified as being in the 'Prodromal' stage of HD. Patients in the Prodromal stage of HD experience irreversible cognitive, motor skill, and functional decline years before the motor onset of HD[1]. A commonly used measure of this

decline is the Composite Unified Huntington's Disease Rating Scale (cUHDRS), a composite score of 4 trackers of motor skill, functional, and cognitive decline which has been proposed for use as a clinical trial continuous measurement endpoint[3].

There are no existing disease-modifying therapeutics for HD, so there is a great need for well-designed clinical trials to evaluate candidate treatments[4]. Historically, clinical trials have tested the efficacy of therapeutics in patients with manifest HD[1]. However, because irreversible neurological damage occurs before motor onset, it is a goal of the field to develop treatments for HD that slow down disease progression in patients in the Prodromal stage of HD[5]. This presents a key challenge. Disease progression is often slow and gradual for many years during the Prodromal stage of HD, accelerating only when a patient is closer to clinical diagnosis[4;6]. Age of clinical diagnosis is highly variable and dependent on patient characteristics such as mutation severity (number of expanded CAG repeats)[7]. A trial recruiting a random sample of participants in the Prodromal stage of HD may show minimal disease progression in the untreated group. These trials would require an unreasonably large sample size and long trial duration to have adequate power to detect a treatment effect if it exists[6].

One strategy to design a well-powered clinical trial with lower sample size and reasonable duration requirements is *sample enrichment*, defined as using patient characteristics to select a study cohort in which there is an increased probability of detecting a treatment effect if it exists[8]. Prognostic sample enrichment strategies increase the probability of detecting a treatment effect by enrolling patients whose health is expected to show a steeper decline if untreated. One proposed prognostic enrichment strategy for clinical trials of Prodromal HD patients is to recruit patients expected to be closer to the motor onset of HD and thus expected to exhibit a faster progression in cUHDRS[5;7;9]. This can be done by recruiting patients using thresholds of "prognostic indicators", proxies for expected time to motor onset which exist in the HD literature. Two of the most commonly used prognostic indicators are "CAP" score and "PIN" score[7;9]. CAP score is a simple function of age and the number of expanded CAG repeats while the PIN score is a function of the CAP score along with a motor function and cognitive performance test (Total Motor Score and Single Digit Modalities Test). The formulas for calculating PIN and CAP can be found in Table 1. A CAP score

of 326 and a PIN score of 0 both correspond to an approximately 50% probability of motor onset of HD within 10 years.

Prognostic Indicator	Formula
CAP	AGE × (CAG-33.66)
PIN	(51 × TMS - 35 × SDMT + 7 × CAP - 883) / 1044

Table 1: Introducing CAP and PIN Scores: TMS = Total Motor Score, SDMT = Single Digit Modality Test, CAG = # of expanded CAG repeats (mutation severity)

Although sample enrichment strategies may increase the feasibility of running a clinical trial, the potential trade-off is the generalizability of trial results to a broader population. This paper will estimate and compare the sample size needed to power a Prodromal HD clinical trial at 90% to detect a treatment effect of a 30% decrease in cUHDRS progression under each of the following enrichment strategies: (1) a CAP score threshold as an inclusion criteria, (2) a PIN score threshold as an inclusion criteria, or (3) no enrichment strategy, i.e. recruiting a random sample of patients in the Prodromal stage of HD with no additional CAP or PIN score inclusion criteria. The demographic and clinical characteristics of the respective eligible participants will be compared to assess the generalizability of trial results under each enrichment strategy to the overall population of patients in the Prodromal stage of HD.

2 Subjects and Methods

2.1 Study Population

ENROLL-HD is a prospective observational study designed to both facilitate research on the natural progression of Huntington's Disease and to act as a platform for clinical trial recruitment[10]. Founded in 2012, ENROLL-HD has recruited over 21 thousand participants - including patients with premanifest, prodromal, and motor onset HD - across 156 clinical sites. Participants undergo annual study visits where they complete clinical tests assessing motor skills, cognition, behaviour, and overall function[10]. Because ENROLL-HD serves as a platform for HD clinical trials recruitment it provides a natural data source for use in sample

size calculations and generalizability assessments.

Data from the ENROLL-HD study collected through October 31, 2020 was used in this paper. ENROLL-HD participants were included in the analyses who were over the age of 18 and not yet diagnosed with motor onset of HD at study baseline, had $\geqslant 36$ CAG repeats, were not missing data necessary to calculate cUHDRS, PIN scores, or CAP scores (Single Digit Modalities Test (SDMT) score, Total Motor (TMS) score, Stroop Word Reading Test (SWR) score or Total Functional Capacity (TFC) score), and had both baseline TMS < 20 and TFC < 11. There were 3,390 eligible participants with a total of 9,068 study visits, a median of 2 study visits, and a median time of 371(interquartile range, 350-411) days between study visits.

2.2 Sample Enrichment

Three separate enrichment strategies were evaluated (1) selecting participants with CAP $\geqslant 336$ ("CAP Enrichment Criteria", (2) selecting participants with PIN $\geqslant -0.1$ ("PIN Enrichment Criteria"), and (3) selecting eligible participants with no additional CAP or PIN criteria (i.e. no enrichment). These CAP and PIN thresholds represent the respective median values of CAP and PIN assessed at baseline in the initial cohort. The full cohort of eligible participants with no enrichment strategy selection criteria had 3,701 participants. 1,678 of these participants met the CAP enrichment criteria and 1,695 of these participants met the PIN enrichment criteria, with an overlap of 1,461 participants.

2.3 Hypothetical Clinical Trial Design

For each of the 3 enrichment strategies (including no enrichment), the sample size was calculated for a hypothetical 2-year-long clinical trial with 5 visits - a baseline visit and 4 follow-up visits spaced every 6 months. A balanced clinical trial sampling schema was used, where each participant had a 50% probability of being assigned to the treatment group. Treatment effect was defined as a reduction in cUHDRS progression and was assumed to be 30%, i.e. a 30% decrease in the magnitude of the time in study slope for cUHDRS.

2.3.1 Outcome Model

The trial was designed so that the treatment effect would be evaluated using the longitudinal linear mixed-effects model

$$
\begin{aligned}
&Y_{ij} = \alpha + \beta \cdot t_{ij} + \gamma \cdot T_i \cdot t_{ij} + a_i + b_i \cdot t_{ij} + \epsilon_{ij} \\
&\epsilon_{ij} \sim N(0, \sigma_e^2), \ \begin{pmatrix} a_i \\ b_i \end{pmatrix} \sim N(0, G) \\
&G = \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{pmatrix}.
\end{aligned} \tag{1}
$$

Y_{ij} refers to cUHDRS for the i^{th} participant ($i = 1, \ldots, N$) measured at their j^{th} visit ($j = 1, \ldots, 5$), α is the fixed population intercept at baseline β is the fixed placebo group slope for cUHDRS, T_i is a treatment indicator for participant i coded as ($T_i = 0$) for patients in the placebo group and ($T_i = 1$) for participants in the treatment group, γ is the fixed population difference in slopes for cUHDRS between the placebo group and the treatment group, and $t_{ij} \in \{0, 0.5, 1, 1.5, 2\}$ is the time in years of visit j for participant i.

a_i is a random individual level deviation about the population intercept, b_i is a random individual level deviation from the population slope, and ϵ_{ij} is random error. Vectors of random effects $\begin{pmatrix} a_i \\ b_i \end{pmatrix}$, $i = 1 \ldots N$, are assumed to be iid. ϵ_{ij}'s are assumed to be iid normally distributed with mean 0 and variance σ_e^2 and independent of the random slopes and intercepts. Visit 1 refers to the baseline visit.

Because the random error and effects are assumed to have mean 0, the expected value of cUHDRS for the treatment and control group can be written respectively as

$$
\begin{aligned}
&E(Y_{ij}|T_{ij} = 0) = \alpha + \beta \cdot t_{ij}, \\
&E(Y_{ij}|T_{ij} = 1) = \alpha + (\beta + \gamma) \cdot t_{ij}.
\end{aligned} \tag{2}
$$

Due to randomization, it was assumed that the population intercept and the variance

components G and σ_e^2 would be the same in the treated and untreated groups.

2.3.2 Hypothesis Test for Treatment Effect

The parameter estimate for γ can be interpreted as the population treatment effect estimate, and testing whether this coefficient is equal to 0 is equivalent to testing for the existence of a treatment effect.

Specifically, the following hypothesis test for a treatment effect was conducted with significance level $\alpha = 0.05$:

$H_0 : \gamma = 0$ versus $H_1 : \gamma \neq 0$.

The test statistic of interest is $t = \frac{\hat{\gamma}}{\sqrt{\widehat{\text{se}}(\hat{\gamma})}}$ which,under H_0, approximately follows a t-distribution with degrees of freedom ($\hat{df}$) estimated using Satterthwaite's approximation[11;12]. This test is implemented in the lmerTest R package[12]. Lower cUHDRS indicates further progression of HD, so $\gamma > 0$ indicates that the treatment had a slowing effect on the progression of HD. Thus a successfully detected beneficial treatment effect is defined as ($|t| \geqslant t_{df,\alpha/2}$ AND $\gamma > 0$).

2.4 Sample Size Calculation by Simulation

A simulation study was conducted using R to estimate the sample size required to reach 90% power to detect a treatment effect of 30% reduction in cUHDRS progression with the Type I error rate controlled at 0.05 for each of the three enrichment strategies. A Monte Carlo approach was used, drawing samples from the model defined in Equation 1 with parameter values determined using estimates from the ENROLL-HD dataset.

2.4.1 Parameter Specification

The three enrichment strategies were used to select three groups of eligible participants from the ENROLL-HD study. For each selected group, parameters for the following linear mixed-effects models were estimated using REML on data from participants with at least 2 study visits:

$$
\begin{aligned}
&Y_{ij} = \alpha + \beta \cdot t_{ij} + a_i + b_i \cdot t_{ij} + \epsilon_{ij} \\
&\epsilon_{ij} \sim N(0, \sigma_e^2), \ \begin{pmatrix} a_i \\ b_i \end{pmatrix} \sim N(0, G) \\
&G = \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{pmatrix}.
\end{aligned} \tag{3}
$$

As in Equation 1 Y_{ij} refers to cUHDRS for the i^{th} participant ($i = 1, \ldots, N$) measured at their j^{th} visit ($j = 1, \ldots, 5$), α is the fixed population intercept at baseline β is the fixed placebo group slope for cUHDRS, T_i is a treatment indicator for participant i, γ is the fixed population difference in slopes for cUHDRS between the placebo group and the treatment group, and t_{ij} is the time in years of visit j for participant i. a_i and b_i represent the random individual level deviation about the population intercept and slope b_i and ϵ_{ij} is random error.

After obtaining parameter estimates $\hat{\alpha}, \hat{\beta}, \hat{\sigma}_e^2$ and $\hat{G}$ the treatment effect parameter $\hat{\gamma}$ corresponding to a 30% treatment effect was calculated as $\hat{\gamma} = -0.3 \cdot \hat{\beta}$. All parameter estimates were plugged into the model defined in Equation 1 to obtain the following model which was used to generate simulated samples:

$$
\begin{aligned}
&Y_{ij} = \hat{\alpha} + \hat{\beta} \cdot t_{ij} + \hat{\gamma} \cdot T_i \cdot t_{ij} + a_i + b_i \cdot t_{ij} + \epsilon_{ij} \\
&\epsilon_{ij} \sim N(0, \hat{\sigma}_e^2), \ \begin{pmatrix} a_i \\ b_i \end{pmatrix} \sim N(0, \hat{G}) \\
&G = \begin{pmatrix} \hat{\sigma}_a^2 & \hat{\sigma}_{ab} \\ \hat{\sigma}_{ab} & \hat{\sigma}_b^2 \end{pmatrix}.
\end{aligned} \tag{4}
$$

2.4.2 Sample Size Scenarios

Power was evaluated for each of the 3 enrichment strategies across a series of sample sizes (total number of trial participants). First, an approximate sample size to reach 90% power, N_{approx}, was calculated using the following closed form sample size estimation equation for

mixed effects models described in chapter 20 of Fitzmaurice, Laird, and Ware[13]

$$N_{approx} = \frac{(Z_{\alpha/2} + Z_{power})^2 \times [\sigma_e^2 \{\tau^2 n(n+1)/12(n-1)\}^{-1} + \sigma_b^2]}{\pi(1-\pi)\gamma^2} \tag{5}$$

where N_{approx} is the approximate sample size, Z_p refers to the pth percentile of a standard normal distribution, τ is the study duration. in years (in this case 2), n is the total number of study visits per participants (in this case 5), and π is the probability of each participant's assignment to the treatment group (in this case 0.5). σ_e^2, σ_b^2, and γ refer to parameters from Equation 1.

Power was estimated via simulation over a grid of sample sizes from $N_{approx} - 200$ to $N_{approx} + 200$ incrementing by 20 for the CAP and PIN enrichment strategies. For the strategy of no enrichment, power was estimated via simulation over a grid of sample sizes from $N_{approx} - 400$ to $N_{approx} + 400$ incrementing by 40 due to the much larger sample size requirement.

2.4.3 Generation of Simulated Data

For each enrichment strategy, 10,000 datasets were simulated to estimate power at each evaluated sample size N. This number was chosen because it led to a Monte Carlo standard error for power of 0.3. Treatment group versus control group membership was assigned to the N participants in each dataset with probability 0.5. Simulated cUHDRS values at baseline and 4 follow-up visits spaced every 6 months were drawn from a distribution under the alternative hypothesis of treatment effect existence, defined by the model described in Equation 4.

Datasets were also generated at each N with cUHDRS values drawn from a distribution defined by the null hypothesis, using the model described in Equation 4 with $\hat{\gamma}$ set to 0.

2.4.4 Power Estimation

The hypothesis test for the existence of a treatment effect described in Section 2.3.2 was conducted for each simulated dataset. Power was estimated as the proportion of datasets simulated under the alternative hypothesis where a treatment effect was successfully de-

tected.

The Monte Carlo standard error of each power estimate was calculated as $SE_P = \sqrt{(\hat{P})(1-\hat{P})/10000}$ where $\hat{P}$ is the power estimated via simulation and 10000 refers to the number of simulation replicates. Approximate sample size ranges to reach 90% power were determined using the lowest and highest N where $\hat{P}$ was within 1 Monte Carlo standard error of 90%.

The Type I Error was estimated as the proportion of datasets simulated under the null hypothesis where a treatment effect was detected.

2.5 Generalizability Analysis

The overlap between the number of ENROLL-HD participants eligible for inclusion under each enrichment strategy was investigated. Demographic and clinical characteristics of the ENROLL-HD participants eligible versus not eligible for recruitment under the CAP and PIN enrichment strategies were compared.

3 Results

3.1 Sample Size Calculation

3.1.1 Parameter Estimation

Parameters were estimated using the model described in Equation 3. All linear mixed effects models converged, and parameter estimates had approximate 95% confidence intervals which did not cross 0. As shown in Table 2, the CAP enrichment strategy led to the steepest estimated slope for cUHDRS of -0.33, indicating that using a CAP threshold for enrichment recruits a cohort with the fastest projected decline in cUHDRS. Using no enrichment strategy led to the smallest estimate of slope magnitude of -0.16. Correspondingly, a 30% treatment effect corresponded to the largest parameter estimate $\hat{\gamma}$ 0f 0.10 for the CAP Enrichement strategy, and the smallest estimate of 0.05 for no enrichment strategy. These parameter estimates were used as settings for the sample size simulation.

	CAP Enrichment	PIN Enrichment	Unenriched
$\hat{\alpha}$ (SE_α)	15.72 (0.06)	15.34 (0.05)	16.39 (0.04)
$\hat{\beta}$ (SE_α)	-0.33 (0.02)	-0.31 (0.02)	-0.16 (0.01)
$\hat{\gamma}$	0.10	0.09	0.05
$\hat{\sigma_e^2}$	0.57	0.59	0.53
$\hat{G}$	$\begin{pmatrix} 3.23 & 0.42 \\ 0.42 & 0.17 \end{pmatrix}$	$\begin{pmatrix} 2.50 & 0.34 \\ 0.34 & 0.20 \end{pmatrix}$	$\begin{pmatrix} 3.12 & 0.37 \\ 0.37 & 0.15 \end{pmatrix}$

Table 2: Parameter Estimates by Enrichment Strategy: The parameter estimates were obtained using a linear mixed effects model to be used as simulation settings for the sample size calculations.

3.1.2 Simulated Sample Size

	CAP Enrichment	PIN Enrichment	Unenriched
N_{approx}	1704	2173	6671
Estimated Sample Size Range	[1643,1683]	[2132,2212]	[6270,6470]

Table 3: Sample Sizes required to reach 90% Power to detect a 30% treatment effect calculated via approximate sample formula (N_{approx}) and simulation (Estimated Sample Size Range)

Table 3 displays the results of the simulated sample size analysis. Figure 2 shows the estimated power evaluated across sample sizes by enrichment strategy, with a Generalized Additive Model used to fit an approximate power curve to the estimates. The CAP enrichment strategy led to the lowest estimated sample size required to reach 90% of between 1624 and 1684. Using no enrichment strategy led to approximately three times the required number of study participants as compared to using an enrichment strategy base on a CAP or PIN threshold.

Type I error was controlled at approximately 0.05 with mean Monte Carlo estimate of type I error of 0.05 with 0.004 standard deviation across all evaluated sample sizes.

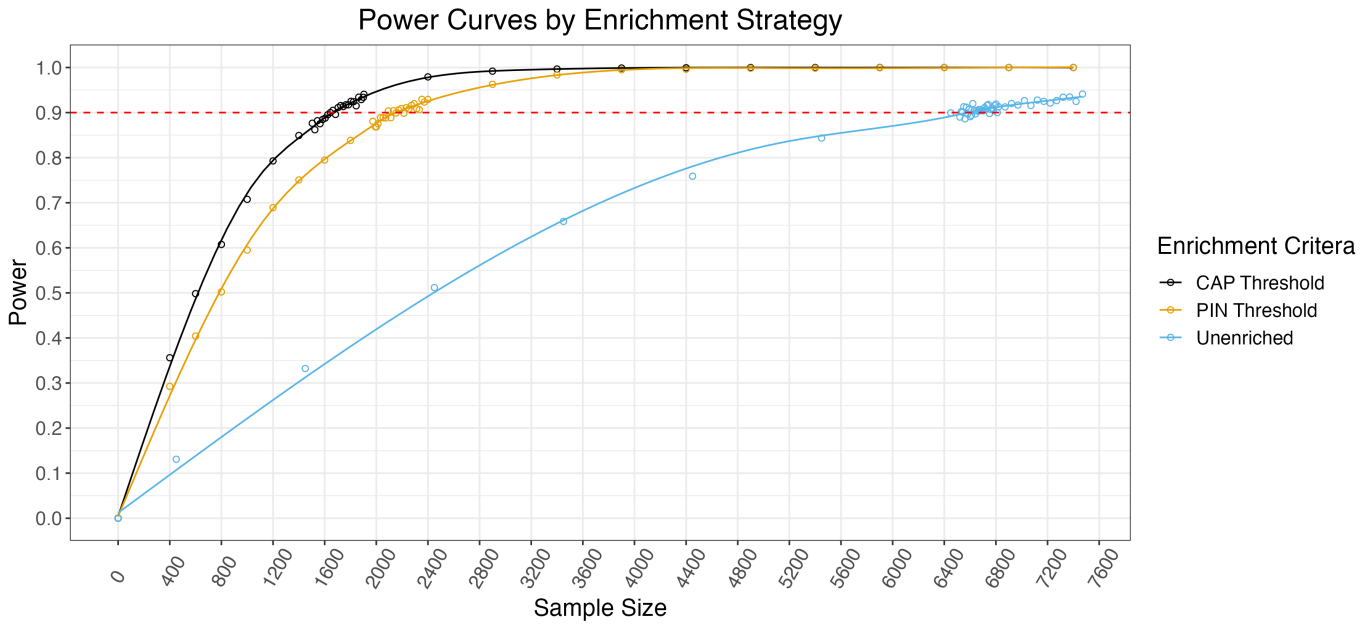

Figure 1: Estimated Power Curves by Enrichment Strategy

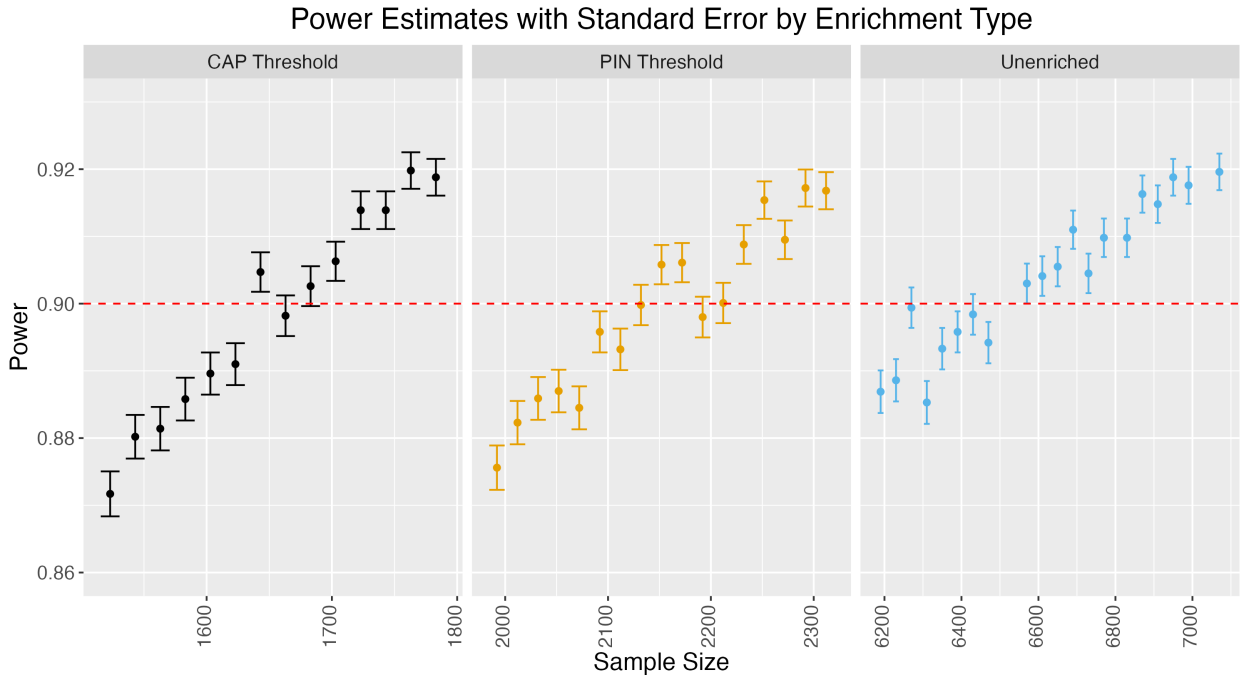

Figure 2: Estimated Power and Monte Carlo SE by Enrichment Strategy

3.2 Generalizability Assessment

As shown in Table 4, ENROLL-HD participants eligible for a clinical trial using the CAP score threshold enrichment strategy were older on average. Eligible participants had in-

creased progression of symptomatic HD as shown by higher baseline Total Motor Scores and lower baseline Stroop Word Reading Test Scores, cUHDRS¡ and Single Digit Modalities Test results.

Similarly, as shown in Table 4, ENROLL-HD participants eligible for a clinical trial using the PIN score threshold enrichment strategy were on average older with increased progression of symptomatic HD as shown by higher baseline Total Motor Scores and lower baseline Stroop Word Reading Test Scores, cUHDRS¡ and Single Digit Modalities Test results.

	PIN Enrichment		**CAP Enrichment**	
	Eligible	Not Eligible	Eligible	Not Eligible
	(N=1695)	(N=1695)	(N=1678)	(N=1712)
Sex				
F	943 (55.6%)	1037 (61.2%)	936 (55.8%)	1044 (61.0%)
M	752 (44.4%)	658 (38.8%)	742 (44.2%)	668 (39.0%)
Age (yrs)				
Mean (SD)	45.1 (12.1)	35.1 (9.95)	44.8 (11.8)	35.4 (10.6)
Race				
Caucasian	1593 (94.0%)	1574 (92.9%)	1587 (94.6%)	1580 (92.3%)
Non-Caucasian	102 (6.0%)	121 (7.1%)	91 (5.4%)	132 (7.7%)
Baseline TMS				
Median [Q1,Q3]	3.00 [1.00,7.00]	0 [0,2.00]	3.00 [0,6.00]	0 [0,2.00]
Baseline SDMT				
Mean (SD)	42.7 (9.82)	56.4 (9.41)	45.5 (11.5)	53.5 (10.8)
Baseline SWR				
Mean (SD)	86.6 (17.9)	98.9 (16.6)	88.7 (18.2)	96.7 (17.5)
Baseline cUHDRS				
Median [Q1,Q3]	15.4 [14.3,16.4]	17.4 [16.5,18.2]	15.7 [14.5,16.9]	17.0 [16.0,18.0]

Table 4: Comparison of Demographic and Baseline Clinical Characteristics of ENROLL-HD participants eligible versus not eligible for a trial using the PIN Enrichment strategy and the CAP Enrichment strategy. TMS = Total Motor Score,and SDMT = Single Digit Modalities Test, SWR - Stroop Word Reading Test score.

1461 participants were eligible under both the PIN and CAP enrichment strategies, whereas 217 were eligible under the CAP but not PIN enrichment strategy and 234 were eligible under the PIN but not CAP enrichment strategy.

4 Discussion

Recruiting eligible participants from the ENROLL-HD study with no enrichment strategy was shown to lead to an astronomically high required sample size of at least 6450 participants to reach 90% power to detect a 30% treatment effect. This requirement would make it financially and logistically unfeasible to run a clinical trial. Both enrichment strategies (inclusion based on of a CAP threshold and inclusion based on a PIN threshold) led to significantly lower required sample sizes. This adds to the body of evidence that HD clinical trials should incorporate a prognostic enrichment strategy to recruit a cohort of participants with faster expected disease progression. Demographic and baseline clinical characteristics differed across the cohort of participants eligible verses not eligible for inclusion in trial when using an enrichment strategy. Enrichment strategy use may make it more difficult to conclude if a treatment is effective in younger patients with more minimal symptomatic progression of HD. Clinicians should consider the trade off between more generalizable trial results and more feasible required numbers of participant. Because there are no existing treatments for HD, it may be more desirable to first run a smaller clinical trial in a cohort of participants with rapid expected disease progression to increase likelihood of detecting a treatment effect. If a treatment effect is successfully detected, further clinical trials could be conducted recruiting from a more general population.

A number of articles have investigated the utility of enrichment strategies for HD trials[6;7;14]. However, those articles have focused on first order sample size approximations to compare enrichment methods or identify optimal prognostic indicator thresholds for enrichment under certain conditions. This paper presents a flexible and more exact method of sample size calculation for HD trials with enrichment. In this paper, sample size was evaluated for a simple longitudinal outcome model with only two fixed and two random effects, the assumption of equality of covariance across the treatment and control groups, and no loss to follow-up. However, this simulation framework could be extended to incorporate additional covariates, allow for unequal covariance matrices across treatment and control groups, and model varying levels of loss to follow-up. A limitation of this analysis is that due to computation limits only 10,000 replications were used per simulation setting. This corresponds to

a Monte Carlo standard error of 0.3 at power = 90%, and led to a level of uncertainty for calculated sample sizes. Future work could rerun this analysis with more replications per simulation setting to reduce noise and give more exact sample size estimates. Uncertainty in the parameter estimates from the model described in Equation 3 was also not addressed into this analysis. Future work could implement further simulation to incorporate the standard error of the parameter estimates and give correspondingly wider error bars on calculated sample sizes.

References

[1] Bates GP, Dorsey R, Gusella JF, et al. Huntington disease (Primer) *Nature Reviews Disease Primers.* 2015.

[2] Wild E.J., Tabrizi S.J.. Premanifest and Early Huntington's Disease in *Huntington's Disease*ch. 5, :86-106Oxford University Press 2014.

[3] 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.

[4] SJ Tabrizi, C Estevez-Fraga, Roon-Mom WMC, et al. Potential disease-modifying therapies for Huntington's disease: lessons learned and future opportunities *Lancet Neurology.* 2022;21.

[5] Paulsen J.S., Langbehn D.R., Stout J.C., et al. Detection of Huntington's disease decades before diagnosis: The Predict HD study *Journal of Neurology, Neurosurgery and Psychiatry.* 2008;79:874–880.

[6] Langbehn DR, Hersch S.. Clinical Outcomes and Selection Criteria for Prodromal Huntington's Disease Trials *Movement Disorders.* 2020;35.

[7] Zhang Y., Long J.D., Mills J.A., et al. Indexing disease progression at study entry with individuals at-risk for Huntington disease *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics.* 2011;156B:751–763. PMCID3174494.

[8] FDA . Enrichment Strategies for Clinical Trials to Support Determination of Effectiveness of Human Drugs and Biological Products - Guidance For Industry tech. rep.Center for Drug Evaluation and Research, Food and Drug Administrationhttps://www.fda.gov/regulatory-information/search-fda-guidance-documents/enrichment-strategies-clinical-trials-support-approval-human-drugs-and-biological-products 2019.

[9] Long J.D., Langbehn D.R., Tabrizi S.J., et al. Validation of a Prognostic Index for Huntington's Disease *Movement Disorders.* 2017;32:256-263.

[10] Sathe S., Ware J., Levey J., et al. Enroll-HD: An Integrated Clinical Research Platform and Worldwide Observational Study for Huntington's Disease *Frontiers in Neurology.* 2021.

[11] Giesbrecht F.G., Burns J.C.. Two-Stage Analysis Based on a Mixed Model: Large-Sample Asymptotic Theory and Small-Sample Simulation Results *Biometrics.* 1985;41.

[12] Kuznetsova A., Brockhoff P., Christensen R.. JSS Journal of Statistical Software December 2017, Volume 82, Issue 13. doi: 10.18637/jss.v082.i13 lmerTest Package: Tests in Linear Mixed Effects Models *Journal of Statistical Software.* 2017;82.

[13] Fitzmaurice M.G., Laird N.M., Ware J.H.. *Applied Longitudinal Analysis.* Wiley 2011.

[14] Paulsen J.S., Lourens S., Kieburtz K., Zhang Y.. Sample enrichment for clinical trials to show delay of onset in huntington disease *Movement Disorders.* 2019.