RESEARCH ARTICLE

Validation of a Prognostic Index for Huntington's Disease

Jeffrey D. Long, PhD,[1,2] Douglas R. Langbehn, MD, PhD,[1,2] Sarah J. Tabrizi, MD, PhD,[3] Bernhard G. Landwehrmeyer, MD, PhD,[4,5] Jane S. Paulsen, PhD,[1,6,7] John Warner, PhD,[4] and Cristina Sampaio, MD, PhD[4,8]

[1]*Departments of Psychiatry and* [2]*Biostatistics, University of Iowa, Iowa City, Iowa, USA*
[3]*Department of Neurodegenerative Disease, UCL Huntington's Disease Centre, University College London Institute of Neurology, London, United Kingdom*
[4]*CHDI Management/CHDI Foundation, Princeton, New Jersey, USA*
[5]*Department of Neurology, University of Ulm, Ulm, Germany*
[6]*Departments of Psychology and* [7]*Neurology, University of Iowa, Iowa City, Iowa, USA*
[8]*Laboratory of Clinical Pharmacology and Therapeutics, University of Lisbon, Lisbon, Portugal*

ABSTRACT: Background: Characterizing progression in Huntington's disease is important for study the natural course and selecting appropriate participants for clinical trials.
Objectives: The aim was to develop a prognostic index for motor diagnosis in Huntington's disease and examine its predictive performance in external observational studies.
Methods: The prediagnosis Neuro-biological Predictors of Huntington's Disease study (N = 945 gene-positive) was used to select a Cox regression model for computing a prognostic index. Cross-validation was used for selecting a model with good internal validity performance using the research sites as natural splits of the data set. Then, the external predictive performance was assessed using prediagnosis data from three additional observational studies, The Cooperative Huntington Observational Research Trial (N = 358), TRACK-HD (N = 118), and REGISTRY (N = 480).
Results: Model selection yielded a prognostic index computed as the weighted combination of the UHDRS total motor score, Symbol Digit Modalities Test, baseline age, and cytosine-adenine-guanine expansion. External predictive performance was very good for the first two of the three studies, with the third being a much more progressed cohort than the other studies. The databases were pooled and a final Cox regression model was estimated. The regression coefficients were scaled to produce the prognostic index for Huntington's disease, and a normed version, which is scaled relative to a 10-year 50% probability of motor diagnosis.
Conclusion: The positive results of this comprehensive validity analysis provide evidence that the prognostic index is generally useful for predicting Huntington's disease progression in terms of risk of future motor diagnosis. The variables for the index are routinely collected in ongoing observational studies and the index can be used to identify cohorts for clinical trial recruitment.

Key Words: Huntington's disease; motor diagnosis; prognostic index; disease progression; external validation

Huntington's disease (HD) is an inherited progressive neurodegenerative disorder caused by a cytosine-adenine-guanine (CAG) repeat expansion in the *HTT* gene. HD signs and symptoms include motor, cognitive, and psychiatric features. There are no current treatments to slow the progression of the disease, but early-phase clinical trials have commenced with the ultimate goal of changing the disease course.[1]

Resources for trial recruitment include clinical research platforms, such as Enroll-HD.[2] Enroll-HD has an embedded observational study with regular visits at which clinical data are collected. Characterizing the progression level of candidate participants is essential for helping to recruit individuals for whom a treatment has measurable efficacy.

Prediagnosis progression has traditionally been characterized by a combination of age and CAG expansion

*Correspondence to: Dr. Jeffrey D. Long, Carver College of Medicine, University of Iowa, 509 Newton Road, Iowa City, IA 52242-1000, USA; E-mail: jeffrey-long@uiowa.edu

Relevant conflicts of interest/financial disclosures: Nothing to report.

Full financial disclosures and author roles may be found in the online version of this article.

Received: 21 July 2016; **Revised:** 6 September 2016; **Accepted:** 9 September 2016

Published online 00 Month 2016 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.26838

because of the well-known association with the timing of motor onset.[3-5] Additional work has shown that the prediction of motor diagnosis, defined as the highest examiner rating on the UHDRS diagnostic confidence level (DCL),[6] is enhanced by considering clinical variables along with the influence of CAG length.[7]

Recently, data from several HD observational studies have become available: PREDICT-HD,[8] COHORT,[9] TRACK-HD,[10] and REGISTRY.[11] The databases provide a unique opportunity to develop a prognostic index for HD (PI_{HD}) and externally validate its predictive performance. Developing a PI_{HD} that can be used across studies would be useful for estimating progression levels in prediagnosis HD and recruiting appropriate candidates for clinical trials. For example, the PI_{HD} can be useful when researchers want to recruit individuals from disparate sources who are just a few years away from estimated diagnosis.

In this study, the prediagnosis PREDICT-HD database will be used to develop a PI_{HD} using CAG, baseline age (age at visit entry), and UHDRS motor and cognitive variables that are common among the studies. Imaging and other technology-intensive variables will be omitted for the pragmatic reasons that they are not measured in all studies and existing registries like Enroll-HD do not have such variables. Once the PI_{HD} is developed with the PREDICT-HD database, the PI_{HD} will be externally validated in the remaining studies by predicting the timing of motor diagnosis in the remaining prediagnosis gene-positive HD individuals.

Patients and Methods

Study Population

Neuro-biological Predictors of Huntington's Disease (PREDICT-HD) is a longitudinal observational study of prediagnosis HD with 32 sites in six countries (AUS, CAN, DEU, ESP, GBR, USA)[8,12-15] with data collected 2002 to 2014. TRACK-HD is a longitudinal prospective observational study of prediagnosis and early HD with four sites in four countries (CAN, FRA, GBR, NE)[10,16,17] with data collected 2008 to 2011. The Cooperative Huntington Observational Research Trial (COHORT) is a longitudinal observational study of HD or at-risk individuals with 38 sites in three countries (AUS, CAN, USA)[9] with data collected 2006 to 2011. REGISTRY is a longitudinal observational study that includes prediagnosis HD, manifest (diagnosed) HD, and at-risk individuals, with a total of 150 mostly European sites[11,18] with data collected 2004 to 2012.

HD gene carriers who did not have a motor diagnosis at baseline (study entry) were used for the analysis. Motor diagnosis was defined as DCL = 4, which is the highest rating and indicates that the examiner has $\geq 99\%$ confidence that the patient exhibits unequivocal HD motor signs.[6] Additional inclusion criteria were ≥ 18 years of age (REGISTRY did not exclude Juvenile HD), a lab-confirmed CAG ≥ 36 (CAG range among studies was 36–66), and complete data on the variables for the analysis. Sample sizes and descriptive statistics for key variables at baseline after applying the inclusion criteria are shown in Table 1.

Study activities were reviewed and approved by institutional review boards (PREDICT-HD, COHORT) or local ethics committees (TRACK-HD, REGISTRY). Informed consent procedures were carried out for each participant, and signed consents for participation and the distribution of deidentified data for collaborative research were obtained.

Statistical Analysis

The first analysis used the PREDICT-HD database for variable selection. Ten preplanned Cox regression models were specified based on previous work with PREDICT-HD.[7] The models were specified to

TABLE 1. Descriptive statistics for variables measured at study entry for the prediagnosis participants meeting the analysis inclusion criteria

	PREDICT	COHORT	TRACK	REGISTRY
N	945	358	118	480
Sites	32	38	4	89
Site sample Size	28.64 (16.42)	9.42 (8.89)	29.50 (0.58)	5.39 (6.02)
Female	605 (64)	225 (63)	64 (54)	257 (54)
Follow-up (years)	4.83 (2.95)	2.06 (1.02)	2.71 (0.64)	2.34 (2.13)
Motor DX	225 (24)	83 (23)	21 (18)	319 (66)
Age	40.40 (10.36)	42.06 (12.51)	40.80 (8.86)	45.19 (12.18)
CAG	42.42 (2.68)	42.41 (2.81)	43.14 (2.41)	43.91 (3.73)
DCL = 3	44 (5)	46 (13)	2 (2)	136 (28)
TMS	4.93 (5.29)	6.20 (8.39)	2.53 (1.68)	12.07 (15.61)
STROOP	98.88 (17.37)	90.62 (19.87)	99.67 (16.10)	79.41 (24.86)
SDMT	50.64 (11.57)	44.93 (12.11)	51.41 (10.24)	38.48 (17.01)

Mean (SD) for quantitative variables and count (percentage) for categorical variables.
Motor DX, prospective motor diagnosis (DCL = 4); TMS, total motor score; STROOP, Stroop word; SDMT, symbol digit modalities test.

represent different levels of parsimony, with the full model (model 1) having the most predictors (see Fig. 1). All models used the predictor scores at baseline to predict the hazard of motor diagnosis (the first occurrence of DCL = 4). A method to reduce bias in model selection is cross-validation (CV), which is often used with several splits of data in order to reduce variance. The participating research sites in PREDICT-HD provided natural and convenient splits (subgroups of participants). Thus, leave-one-site-out CV (LOSO-CV) was used for model selection.[19] For LOSO-CV, each Cox model was estimated using all the sites but one, and then the performance of the model for predicting the timing of motor diagnosis was evaluated in the omitted site. The process was repeated leaving each site out in turn, resulting in a predictive performance index for all 32 sites. For each estimated Cox model, the PI_{HD} was computed as the weighted combination of the regression weights and the predictor scores. Focus was on prediction, so the proportional hazards assumption was not a primary concern. However, Schoenfeld residuals showed general consistency with the proportional hazards assumption for the main model discussed below.

Predictive performance was assessed using the integrated Brier score,[20] Harrell's *C* (concordance) statistic,[21] and a *C* statistic weighted for censoring.[22] The indexes provided similar model evaluations, and we focus on Harrell's *C* because it is the most commonly reported (extended results are presented in the appendix). *C* indexes the ability of the PI_{HD} to discriminate between individuals with longer time to motor diagnosis and those with shorter time. *C* is similar to the area under the receiver operating characteristic curve in sensitivity/specificity analysis, but it accounts for the right-censoring attributing to dropout and study termination. $C = 0.5$ indicates no better than chance prediction of the timing of motor diagnosis, whereas $C = 1$ is perfect prediction (each pair of individuals is correctly ordered in terms of timing of diagnosis). A guideline for the evaluation of the magnitude of *C* is provided by a survey of values found in external validation studies in oncology and cardiovascular disease.[23] The survey showed mean $C=0.78$, and we consider this value to be the benchmark for typical and acceptable predictive performance.

The second analysis was an external validation that included all the databases and used the model selected in the first analysis. To provide a benchmark for comparison, internal validation for each study was performed using LOSO-CV, and the mean *C* was computed among sites. For the external validation, the weights estimated from PREDICT-HD were used to compute the PI_{HD} in each other study, then *C* was computed. The Cox regression weights and standard errors (SEs) of the PI_{HD} predictors were estimated independently in each study (values were multiplied by 1,000 to avoid very small numbers). A significance test of the difference between a study's regression weight and the corresponding PREDICT-HD regression weight was computed using a multiple-group analysis based on dummy coding with combined databases.[24]

Risk groups for motor diagnosis were formed by using the quartiles of the PI_{HD} as computed in PREDICT-HD applied to all the studies. In PREDICT-HD, the quartiles divided the PI_{HD} distribution into four equal groups (25% each). These same quartiles were used in the other studies, but they did not necessarily divide the other distributions into quarters. Survival was defined as the probability of not having a motor diagnosis, and survival curves were estimated using cubic splines[25] to provide smooth curves over the follow-up period. Similar cubic spline survival curves and bootstrapped 95% confidence intervals were constructed for the pooled data as described below.

Results

Model selection results using only the PREDICT-HD database are shown in Figure 1. The figure consists of box plots of distributions of Harrell's *C* among the sites for the 10 models. Predictive performance varied by site, with some sites showing worse than random performance for some models ($C < 0.5$) and other sites showing perfect performance ($C = 1.0$). Several models had similar good performance (model 1 through model 6), but the summary statistics slightly favored model 4. Model 4 had the largest median ($C = 0.87$) and the largest lower and upper quartiles (25% = 0.82; 75% = 0.93), indicating its *C* distribution had the closest proximity to perfect prediction (shifted closest to the top of Fig. 1). Similar results were shown for the Brier score and the weighted *C*, and model 4 had the best aggregate rank among all the measures (see appendix). Therefore, we considered model 4 to be an adequate model to carry forward for further consideration (other models could have been carried forward as well). In all subsequent analysis, the PI_{HD} was computed as the weighted composite of the UHDRS total motor score (TMS), Symbol Digit Modalities Test (SDMT), and the CAG-Age Product (CAP),[5] with the latter being CAP = Age at baseline$\times$(CAG$-$34).

External validation results using *C* are shown in Table 2. For COHORT and TRACK-HD, the PI_{HD} external predictive performance (C_{EXT}) was very similar to the internal performance (C_{INT}). For example, COHORT had internal performance of $C_{INT}=0.85$ and external performance of $C_{EXT}=0.84$. The REGISTRY study showed low internal predictive performance ($C_{INT}=0.64$) and even lower external performance ($C_{EXT}=0.56$). The Cox regression weights for TRACK were larger in absolute value than their PREDICT-HD counterparts and much more

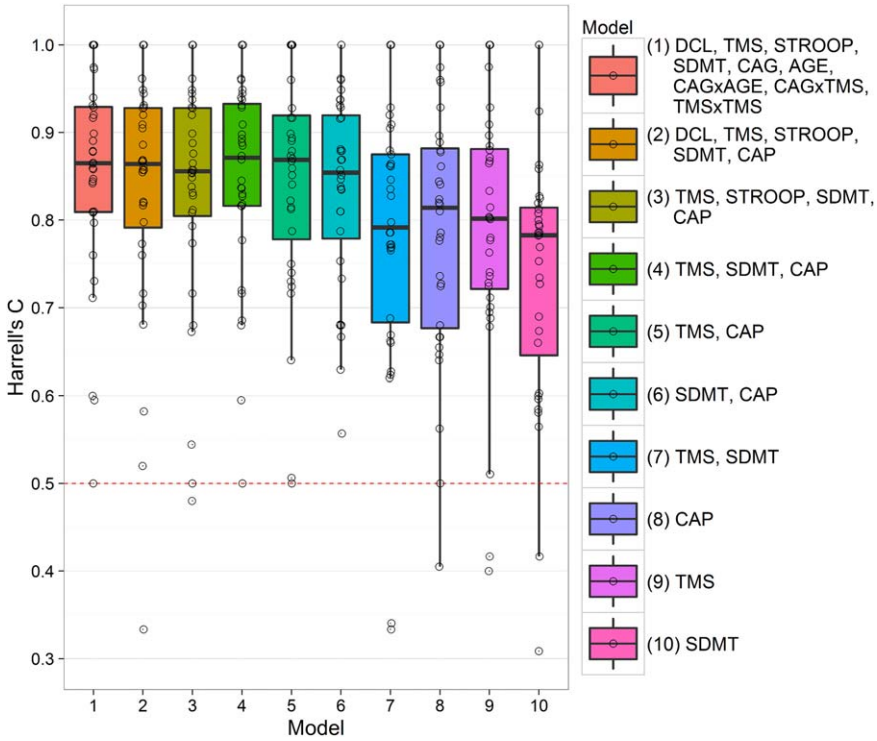

FIG. 1. Boxplots and individual values (circles) of Harrell's *C* for 10 models among the PREDICT sites. TMS, total motor score; STROOP, Stroop word; SDMT, symbol digit modalities test; AGE, baseline age; CAP = Age x (CAG - 34).

so for TMS. However, the estimated SEs were also large and there were no statistically significant differences. (Entry criteria for the TRACK-HD study included a maximum screening TMS of 5, and the restricted range may be related to these estimates.) The COHORT SDMT weight was significantly smaller than the PREDICT-HD weight, but the differences for TMS (larger) and CAP (smaller) were not significant. Each REGISTRY regression weight was significantly smaller in absolute value than the corresponding PREDICT-HD weight, which was indicative of the general weakness of the predictors in the REGISTRY database.

Figure 2 shows the smooth survival curves for the motor diagnosis risk groups along with the counts and percentages of group membership (survival is the probability of not being diagnosed). The risk gradient of the survival curves displayed by PREDICT-HD (upper left panel) was similar for COHORT (upper right) and TRACK-HD (lower left) for the early years (up to year 3) in the sense that the risk curves were in the same order from top to bottom (low $\geq$ mid-low $>$ mid-high $>$ high). In contrast, the curve ordering for REGISTRY (lower right) was inconsistent, given that there were several order changes (crossing lines) in the early years among the three lower-risk groups. However, similar to the other studies, the REGISTRY high-risk group curve was consistently the lowest over time (there was some early overlapping with mid-high). By design, the PI_{HD} quartiles define four equal groups for PREDICT-HD, but the same PREDICT-HD PI_{HD} cutoffs did not result in equal groups in the other studies. However, COHORT and TRACK-HD had a similar 50% split as PREDICT-HD into lower risk of motor diagnosis (low + mid-low) and higher risk of motor diagnosis (mid-high + high). On the other hand, REGISTRY had more skewed classification percentages,

TABLE 2. External validation results: *C* statistic (internal and external validation), Cox regression weights (SEs), and tests of difference with the PREDICT weights

Study	C_{INT}	C_{EXT}	TMS*	SDMT[a]	CAP[a]
PREDICT	0.85		56.93 (9.44)	−35.01 (7.02)	7.63 (0.79)
COHORT	0.85	0.84	58.25 (12.71)	−22.45*** (13.09)	4.46 (1.41)
TRACK	0.79	0.79	219.99 (121.86)	−61.45 (27.21)	9.23 (5.39)
REGISTRY	0.64	0.56	0.43*** (5.27)	−5.56*** (5.05)	2.33** (0.45)

[a]Regression weights and SEs are multiplied by 1,000 to avoid small numbers; C_{INT} = internal validation Harrell's *C;* C_{EXT} = external validation; CAP = Age × (CAG − 34).
***$P < 0.001$; **$P < 0.01$; *$P < 0.05$.

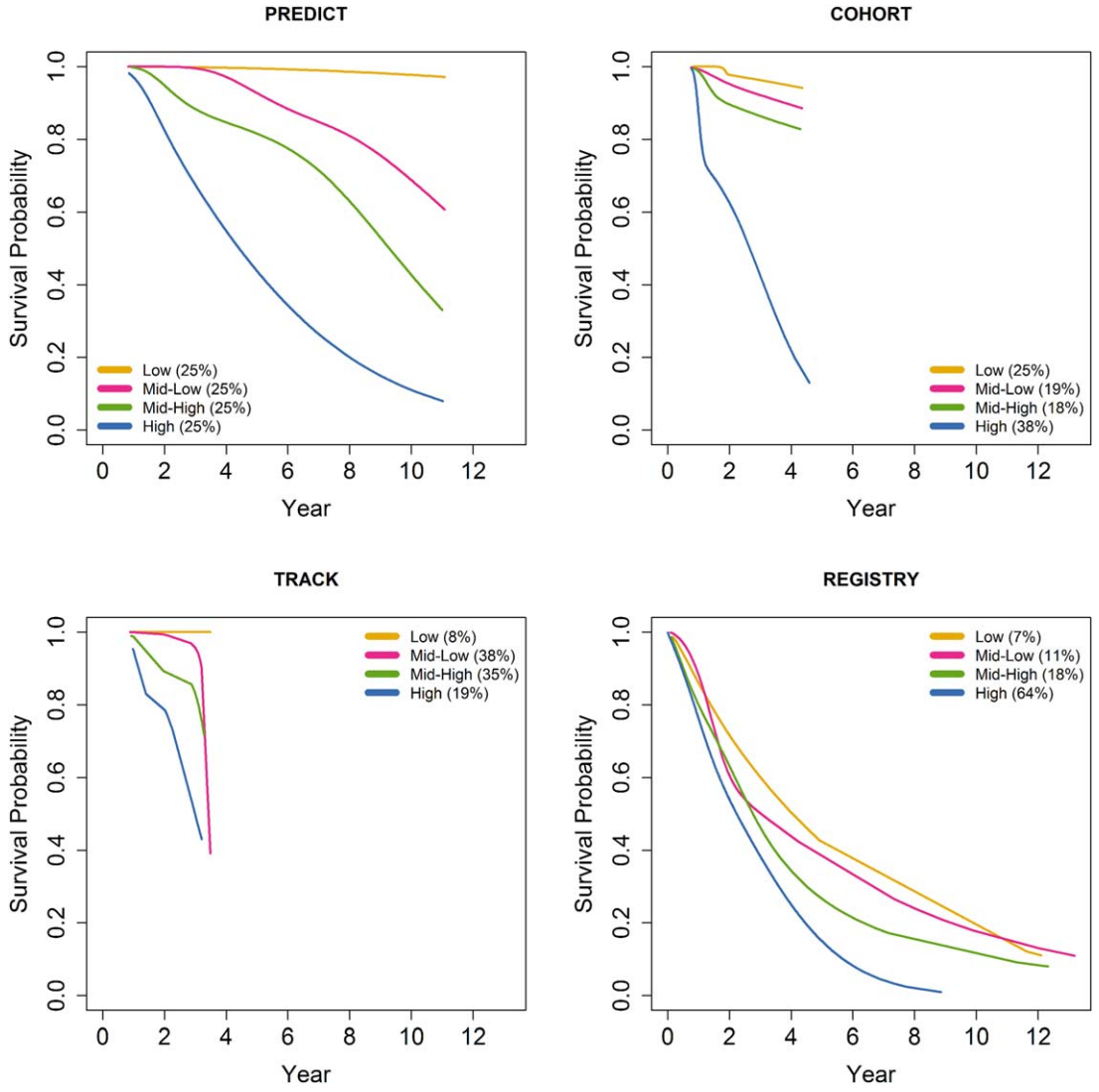

FIG. 2. Cubic spline survival curves for risk groups of motor diagnosis defined by the PREDICT-HD study cutoffs. Counts (percentages) are shown for each group.

with 82% of the sample classified as having a higher risk of motor diagnosis.

Pooling among databases is justified when baseline hazards (and survival curves) are similar.[26] Cubic spline estimates of the baseline hazards are shown in the appendix. Similar to Figure 2, there was reasonable similarity among PREDICT-HD, COHORT, and TRACK-HD. Thus, the three databases were combined, and a final Cox regression model was estimated. Based on the estimated Cox regression weights (multiplied by 1,000 and rounded for simplicity), the formula is computed as

$$\mathrm{PI}_{HD} = 51 \times TMS + (-34) \times \mathrm{SDMT} + 7 \times \mathrm{Age} \times (\mathrm{CAG} - 34).$$

Larger values index greater risk of motor diagnosis and hence greater predicted progression. For example, if a person has TMS = 10, SDMT = 90, age = 40, and CAG = 42, then $\mathrm{PI}_{HD} = -310$, whereas a person with TMS = 15, SDMT = 70, age = 50, and CAG = 42 has $\mathrm{PI}_{HD} = 1{,}185$. The PI_{HD} was "normalized" (scaled) to enhance interpretation. The survival curve was estimated for each PI_{HD} based on the fitted model. For the 10-year time point, the PI_{HD} value associated with a 0.5 survival probability was selected as the centering constant (=883), and the standard deviation (SD) of the PI_{HD} distribution was selected as the scaling constant (=1,044). The prognostic index normed for HD (PIN_{HD}) is

$$PIN_{HD} = \frac{\mathrm{PI}_{HD} - 883}{1044}.$$

PIN_{HD} is in SD units and is interpreted relative to 50% 10-year survival. $PIN_{HD} < 0$ indicates greater than 50% 10-year survival, and $PIN_{HD} > 0$ indicates less than 50% 10-year survival. From the example above, the person with $\mathrm{PI}_{HD} = -310$ has $\mathrm{PIN}_{HD} = -1.14$ that is 1.14 SDs below the 50% 10-year survival mark, and the person with $\mathrm{PI}_{HD} = 1185$ has $\mathrm{PIN}_{HD} = 0.29$ that is

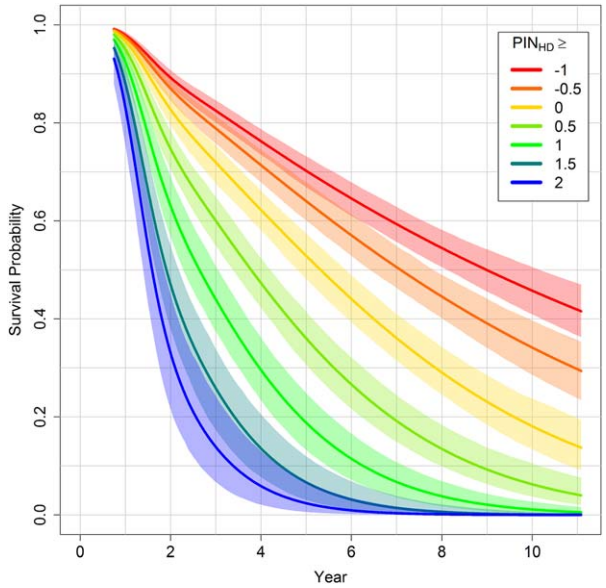

FIG. 3. Cubic spline survival curves (95% CIs) for ranges of the prognositc index normed for Huntington's disease (PIN_{HD}). Curves are based on pooling PREDICT-HD, TRACK-HD, and COHORT.

0.29 SDs above the mark. Figure 3 shows the survival curves and bootstrapped 95% CIs for different ranges of PIN_{HD} defined by a lower cut-off value. For example, a lower cutoff of 1 defines the right-tail range of $PIN_{HD} \geq 1$ for the pooled data. For this cohort, the figure indicates that 40% are expected to have a motor diagnosis a little after 2 years (60% remain prediagnosis), and 70% are expected to have a motor diagnosis by 4 years (30% remain prediagnosis).

Discussion

To our knowledge, this is the first external validation of a prognostic index for motor diagnosis (conversion) in prediagnosis HD. External validation is the gold standard for reproducibility,[27,28] and the very good predictive performance in two out of the three external studies underscores the general usefulness of the PI_{HD}. The external predictive performance was slightly higher than that found in similar validation studies.[23]

The poor performance in the REGISTRY study can be explained by a lack of variability in progression level. The REGISTRY cohort was much more progressed at baseline, with many individuals having clinical scores in the ranges typically associated with diagnosis. The lack of progression variability resulted in the PI_{HD} variables (TMS, SDMT, and CAP) having limited predictive power. Like the other studies, REGISTRY had a standard protocol, but it was the only study to enroll individuals less than 18 years of age, it had the largest variety of languages (13), the largest number of sites (see Table 1), and all the sites were European for our analysis.

Though the predictive performance of the PI_{HD} was poor for the REGISTRY study, this does not necessary threaten the validity of the PI_{HD}. Consistent with conventional benchmarks of motor and cognitive performance, the PI_{HD} was successful in correctly classifying most of the REGISTRY individuals as having a high risk of motor diagnosis.

The PI_{HD} or its normed version (PIN_{HD}) can be used to predict progression, with higher scores indicating greater risk of motor diagnosis. For clinical purposes, predicted progression provides general information about the current status and course of the disease. The survival curves indicate what happens to an individual's cohort with a common PI_{HD}. Such information can be useful for treatment strategies and life planning.

For research purposes, PI_{HD} provides a more accurate index. Progression is commonly indexed by the CAP, which is a type of "burden score" of age and CAG expansion that has several variants.[3-5,29] Our results show that when CAP is supplemented with the TMS and SDMT, predictive performance increases. The added variables include one of the core components of the UHDRS, the summation of motor signs (TMS), and the other is one of the three UHDRS cognitive tests (SDMT) that also has a motor component.

Because PIN_{HD} is easy to compute with commonly measured variables, it is useful for recruiting cohorts of individuals for clinical trials from among those who are registered in ongoing observational studies, such as Enroll-HD. The anonymized Enroll-HD database is available to researchers online (https://www.enroll-hd.org/for-researchers/access-data/), and the PIN_{HD} can be computed using an individual's last wave of data. There is an optional consent for participation in future research studies, which has a very high rate of endorsement (>90%).[2] One use of PIN_{HD} is to identify consenting individuals who have progression levels appropriate for a clinical trial. If a preventative trial requires participants to be many years from motor diagnosis, then a certain bottom percent (e.g., 20%) of individuals with the smallest PIN_{HD} values might be selected from Enroll-HD. More finely tuned selection is possible if the researcher can identify the desired proportion of converters at the end of the proposed study. Each PIN_{HD} value (or range) is associated with a cohort that has a specific expected rate of conversion over a given time interval. After deciding upon the nontreated survival curve that is appropriate for the clinical trial, the potential pool of candidates can be identified by the associated PIN_{HD} value(s). For example, an intervention that targets relatively advanced progression might require the untreated group to have 40% survival (60% conversion) at a study terminus of 3.5 years. Figure 2 indicates that the appropriate survival curve is associated with the range of $PIN_{HD} \geq 1$. Consenting individuals from Enroll-

HD whose PIN_{HD} falls within this range might be candidates for recruitment.

In addition to identifying whom to recruit, PIN_{HD} is useful for determining the sample size when the endpoint of the clinical trial is conversion itself (time to DCL = 4). In such trials, the goal is to determine whether motor diagnosis is delayed by treatment relative to placebo. The expected number of events and the hypothesized treatment effect on the hazard ratio are the key determinants of sample size for survival-based clinical trials.[30] The survival curve for a PIN_{HD} can be used to estimate the number of expected conversions in the (untreated) placebo group at study end.

PIN_{HD} can also be informative for recruitment when the trial is not directly concerned with survival, as when change in TMS is the outcome. Based on the analysis of large, observational studies, it is well established that the rate of deterioration for imaging and clinical variables accelerates as they increasingly depart from control-like initial levels.[10,15,17,31] The rate of deterioration also increases with CAP.[32] Because PIN_{HD} is based on clinical variables and CAP, deterioration of clinically relevant outcome measures is expected to accelerate as PIN_{HD} increases. Consider the example of change in intracranial volume–corrected putamen volume $\times 1,000$ (the scaling is used to avoid small numbers) using the PREDICT-HD database. We found that $PIN_{HD} = -1$ was associated with an annual rate of decrease of $\beta = -0.0627$ (95% CI = [–0.0714, –0.0539]), and $PIN_{HD} = 1$ was associated wth the greater rate of $\beta = -0.0881$ (95% CI = [–0.0987, –0.0774]).

An ancillary analysis with results presented in the appendix showed that the weights of the PI_{HD} and its predictive performance were similar when estimated separately at seven annual follow-up visits. Thus, there is evidence that the PI_{HD} will not change much when computed over a several-year span. The implication for recruitment is that if individuals have repeated observations, then only the last wave of data might be sufficient for predicting progression level.

The emphasis of this study was on developing a prognostic index with variables that are routinely collected in observational studies. Potentially important predictors, such as putamen volume, were intentionally excluded because COHORT and REGISTRY did not have imaging, and neither does Enroll-HD. Recent results using PREDICT-HD[7] show that imaging variables were predictive of motor diagnosis when considered in isolation. However, when important clinical variables (e.g., TMS, SDMT) and genetic variables (i.e., CAP) were already in the model, the boost in predictive power was modest when imaging variables were added. Therefore, the exclusion of imaging variables does not substantially lower the predictive power of the PI_{HD}. A potentially more powerful prognostic index could be developed using a combination of newly developed or discovered imaging, clinical, genetic, or wet biomarker variables.

A caveat regarding the analysis is that only individuals without a motor diagnosis at baseline were analyzed. This resulted in the exclusion of diagnosed participants. There is a potential for selection bias if the included and excluded individuals were representative of different populations.

The databases used in this study may not represent the more general clinic population. PREDICT-HD and TRACK included participants who had premanifest genetic testing. Individuals who have such testing may be less than 10% of the HD population in North America[31] (though the percentage is probably higher in Europe and elsewhere).

Another qualification is the overlap of participants for COHORT and PREDICT-HD. The data from these studies were concurrently collected and they had many sites in common. Because of patient confidentiality and the anonymized data, it is not possible to identify unique participants. Nonetheless, we did use birth year, CAG length, and sex to examine commonalities among the studies. Approximately one third of the COHORT subjects could not be mapped to the PREDICT-HD subjects (overlap is probably 75% at most).

We caution that the PIN_{HD} computed for one person does not necessarily predict progression for that individual. The survival curves of the models and the figures indicate the risk of motor diagnosis for a cohort with a common PIN_{HD}. Because of individual variability, prediction for a single person is difficult, and the survival curve for a person's cohort may have little bearing on their actual risk for motor diagnosis.[32]

Finally, our findings demonstrate that a relatively simple risk score computed on readily available variables measured at a single visit has reasonable generalizability and usefulness for predicting progression toward HD motor diagnosis, and thus may be valuable for targeted recruitment for clinical trials. Further analysis will be needed to assess whether usefulness is improved adding newly developed variables. ■

Acknowledgments: Samples and/or data from the COHORT Study, which received support from HP Therapeutics, Inc., were used in this study. We thank the Huntington Study Group COHORT investigators and coordinators who collected data and/or samples used in this study, as well as participants and their families, who made this work possible. TRACK-HD was supported by CHDI Foundation, a nonprofit biomedical research organization dedicated to developing therapeutics for Huntington's disease. We thank the TRACK-HD study participants and their families. We also acknowledge the support of the National Institute for Health Research University College London Hospitals Biomedical Research Centre and the Manchester Biomedical Research Centre. PREDICT-HD was supported by the National Institutes of Health under the following grants: 5R01NS040068, 5R01NS054893, 1S10RR023392, 1U01NS082085, 5R01NS050568, 1U01NS082083, and 2UL1TR000442-06 (J.S. Paulsen principal investigator). This research was also supported by CHDI Foundation grant A3917 and the National Alliance for Medical Image Computing, which provided general data collection/analysis support. We thank the staff at the PREDICT-HD sites, the study participants, the National

Research Roster for Huntington Disease Patients and Families, the Huntington's Disease Society of America, and the Huntington Study Group. We thank the REGISTRY participants and families, CHDI, EHDN, and the REGISTRY investigators of EHDN.

References

1. Wild EJ, Tabrizi SJ. Targets for future clinical trials in Huntington's disease: What's in the pipeline. Mov Disord 2014;29:1434-1445.
2. Landwehrmeyer BG, Fitter-Attas C, Giuliano J, et al. Data analytics from Enroll-HD, a global clinical research platform for Huntington's disease. Mov Disord Clin Pract 2016 Jun 22. doi: 10.1002/mdc3.12388. [Epub ahead of print]
3. Langbehn DR, Hayden MR, Paulsen JS. 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 2010;153: 397-408.
4. Penney JB, Vonsattel JP, MacDonald ME, Gusella JF, Myers RH. CAG repeat number governs the development rate of pathology in Huntington's disease. Ann Neurol 1997;41:689-692.
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 B Neuropsychiatr Genet 2011;156:751-763.
6. Huntington Study Group. Unified Huntington's Disease Rating Scale: reliability and-consistency. Mov Disord 1996;11:136-142.
7. Long JD, Paulsen JS. Multivariate prediction of motor diagnosis in Huntington disease: 12 years of PREDICT-HD. Mov Disord 2015; 12:1664-1672.
8. Paulsen J, Long J, Ross C, et al. Prediction of manifest Huntington's disease with clinical and imaging measures: a prospective observational study. Lancet Neurol 2014;13:1193-1201.
9. Dorsey ER; the Huntington Study Group COHORT Investigators. Characterization of a large group of individuals with huntington disease and their relatives enrolled in the COHORT study. PLoS One 2012;7:e29522.
10. Tabrizi SJ, Scahill RI, Owen G, 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 2013;12:637-649.
11. Handley O, Landwehrmeyer B; REGISTRY Steering Committee and the EHDN REGISTRY Investigators. European Huntington's Disease Network REGISTRY: current status. J Neurol Neurosurg Psychiatry 2012;83:A47.
12. Paulsen JS, Hayden M, Stout JC, et al. Preparing for preventive clinical trials: the Predict-HD study. Arch Neurol 2006;63:883-890.
13. Paulsen JS, Langbehn DR, Stout JC, et al. Detection of Huntington's disease decades before diagnosis: the PREDICT-HD study. J Neurol Neurosurg Psychiatry 2008;79:874-880.
14. Paulsen JS, Wang C, Duff K, et al. Challenges assessing clinical endpoints in early Huntington disease. Mov Disord 2010;15:2595-2603.
15. Paulsen JS, Long JD, Johnson HJ, et al.; PREDICT-HD Investigators and Coordinators of the Huntington Study Group. Clinical and biomarker changes in premanifest Huntington disease show trial feasibility: a decade of the PREDICT-HD study. Front Aging Neurosci 2014;6:78.
16. Tabrizi SJ, Langbehn DR, Leavitt BR, et al. Biological and clinical manifestations of Huntington's disease in the longitudinal TRACK-HD study: cross-sectional analysis of baseline data. Lancet Neurol 2009;8:791-801.
17. Tabrizi S, Scahill R, Durr A, et al. Biological and clinical changes in premanifest and early stage Huntington's disease in the TRACK-HD study: the 12-month longitudinal analysis. Lancet Neurol 2011;10:31-42.
18. Orth M, Handley OJ, Schwenke C, Landwehrmeyer B. Observing Huntington's disease: The European Huntington's Disease Network's REGISTRY. J Neurol Neurosurg Psychiatry 2010;82:1409-1412.
19. König IR, Malley JD, Weimar C, Diener HC, Ziegler A. Practical experiences on the necessity of external validation. Stat Med 2007; 26:5499-5511.
20. Gerds TA, Schumacher M. Consistent estimation of the expected Brier score in general survival models with right-censored event times. Biom J 2006;48:1029-1040.
21. Harrell FE, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. J Am Med Assoc 1982;247:2543-2546.
22. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei WL. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. Stat Med 2011;30:1105-1117.
23. Collins GS, Groot JA de, Dutton S, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. BMC Med Res Methodol 2014;14: 40-51.
24. Clogg CC, Petkova E, Haritou A. Statistical methods for comparing regression coefficients between models. Am J Sociol 1995;100: 1261-1312.
25. Royston P, Parmar MK. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. BMC Med Res Methodol 2013;13:152-162.
26. Royston P, Altman DG. External validation of a cox prognostic model: principles and methods. BMC Med Res Methodol 2013;13: 33-48.
27. Bouwmeester W, Zuithoff NP, Mallett S, et al. Reporting and methods in clinical prediction research: a systematic review. PLoS Med 2012;9:1-12.
28. Debray TP, Vergouweb Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KG. A new framework to enhance the interpretation of external validation studies of clinical prediction models. J Clin Epidemiol 2015;68:279-289.
29. Ross CA, Aylward EH, Wild EJ, et al. Huntington disease: natural history, biomarkers and prospects for therapeutics. Nat Rev Neurol 2014;10:204-216.
30. Machin D, Campbell MJ, Tan SB, Tan SH. Sample Size Tables for Clinical Studies. Hoboken, NJ: Wiley-Blackwell; 2009.
31. Paulsen JS, Smith MM, Long JD. Cognitive decline in prodromal Huntington disease: implications for clinical trials. J Neurol Neurosurg Psychiatry 2013;84:1233-1239.
32. Long JD, Paulsen JS, Marder K, Zhang Y, Kim J, Mills JA. Tracking motor impairments in the progression of Huntington's disease. Mov Disord 2014;29:311-319.
33. Shoulson I, Young AB. Milestones in Huntington's disease. Mov Disord 2011;26:1127-1133.
34. Henderson R, Keiding N. Individual survival time prediction using statistical models. J Med Ethics 2005;31:703-706.
35. Kolde R, Laur S, Adler P, Vilo J. Robust rank regression for gene list integration and meta-analysis. Bioinformatics 2012;28:573-580.

Supporting Data

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.