Plasma NfL as a prognostic biomarker for enriching HD-ISS stage 1 categorisation: a cross-sectional study

Georgia M. Parkin,[a,c,] Elizabeth A. Thomas,[b,c] and Jody Corey-Bloom[a]*

[a]Department of Neurosciences, University of California San Diego, San Diego, CA 92093, USA
[b]Department of Neurobiology and Behavior, University of California Irvine, Irvine, CA 92697, USA
[c]Institute for Interdisciplinary Salivary Bioscience Research, University of California Irvine, Irvine, CA 92697, USA

Summary

Background The recently proposed Huntington's Disease Integrated Staging System (HD-ISS) categorises individuals with the *Huntintin* genetic mutation into disease progression cohorts based on quantitative neuroimaging, cognitive, and functional markers for research purposes. Unfortunately, many research studies do not collect quantitative neuroimaging data, and so the authors of the HD-ISS have subsequently provided approximated cohort thresholds based on disease and clinical data alone. However, these are rough proxies that aim to maximise stage separation, and should not be considered as 1:1 substitutes for the HD-ISS. Notably, no wet biomarker met the stringent criteria required to be considered a landmark for HD-ISS categorisation. We have previously shown that levels of plasma neurofilament light (NfL), a neuronal marker associated with axonal injury, are associated with predicted years to clinical motor diagnosis (CMD). Our objective in the current study was to determine whether HD-ISS categorisation, particularly for stages prior to CMD, could be improved with consideration of plasma NfL levels.

Methods A total of 290 blood samples, and clinical measures, were collected from participants across all HD-ISS stages: n = 50 [Stage 0], n = 64 [Stage 1], n = 63 [Stage 2], n = 63 [Stage 3], as well as 50 healthy controls. Plasma NfL levels were measured using a Meso Scale Discovery assay.

Findings Cohorts differed by age, cognitive function, CAG repeat length, and select UHDRS measures. Plasma NfL levels also differed significantly across cohorts. Approximately 50% of Stage 1 participants had plasma NfL levels indicative of predicted CMD within ten years.

Interpretation Our findings suggest that plasma NfL levels may have use in enriching Stage 1 membership into subgroups that are less than, and within, predicted 10 years until CMD.

Keywords: Huntington's disease; Biomarkers; Neurofilament light; Plasma

Introduction

Huntington's Disease (HD) is a genetic, neurodegenerative disorder caused by an autosomal dominant mutation in the *Huntingtin (HTT)* gene, which results in an expansion of a trinucleotide (CAG) repeats in exon one. Specifically, the normal number of HTT gene CAG repeats is <35, individuals with partial penetrance for HD have 36–39 repeats, and those predisposed to develop HD have 40+ repeats. For those who carry the genetic mutation, which can be confirmed via a genetic test, the hereditary nature of the mutation confers a certainty of symptom onset within one's lifetime. Symptom onset occurs most often during an individual's third to fifth decade of life, and commonly begins with sub-clinical or premanifest behavioural and mood disturbances, followed by insidious decline. A clinical motor diagnosis (CMD) is provided following the onset of motoric symptoms unequivocally associated with HD, with deterioration ultimately leading to a premature death around 20 years after diagnosis. Disease severity increases with CAG repeat length, and as such age and CAG repeat length can be used to estimate

*Corresponding author. Department of Neurosciences, University of California, San Diego, La Jolla, CA, USA.
    E-mail address: gparkin@health.ucsd.edu (G.M. Parkin).

Research in context

Evidence before this study

Huntington's Disease (HD) is a genetic, neurodegenerative disorder caused by an autosomal dominant mutation in the *Huntingtin (HTT)* gene. Symptom progression is insidious, beginning in the third to fifth decade of life and ultimately resulting in a premature death 20 years after motor symptom onset and corresponding clinical motor diagnosis (CMD). Recently, researchers Tabrizi and colleagues proposed the HD Integrated Staging System (HD-ISS), which categorises individuals with the *HTT* genetic mutation into disease progression cohorts based on quantitative neuroimaging, cognitive, and functional markers, for research purposes. The HD-ISS has also been adopted into some clinical trial selection criteria. Currently, no wet biomarker has met the stringent criteria required to be considered a landmark in the HD-ISS. We have previously shown that plasma levels of neurofilament light (NfL), a neuronal marker associated with neurodegeneration, is associated with predicted years to CMD and proposed that a plasma NfL cut-point of ≥45.0 pg/ml could distinguish participants who are within ten years of CMD.

Added value of this study

This is the first study to assess whether consideration of plasma NfL levels may assist in enriching stage membership in the newly proposed HD-ISS, particularly for stages that precede CMD (Stage 3). We show that approximately 50% of HD-ISS Stage 1 participants had plasma NfL levels indicative of predicted CMD within ten years.

Implications of all the available evidence

We contend that plasma NfL levels may have use in enriching Stage 1 membership into sub-groups that are less than, and within, predicted 10 years until CMD. Such enrichment will allow for more accurate analysis of early HD disease progression cohorts.

lifetime exposure to the *HTT* mutation, or disease burden, and predict age of or years to CMD.[1,2] Unfortunately, use of these computations is currently confined to research purposes, as their accuracy is limited and variability is high. It has been estimated that up to 40% of the variability in age of CMD is governed by genes exclusive of *HTT*, and also influenced by environmental factors.[1,3,4] The ability to accurately predict CMD and disease progression is clinically meaningful, as it would allow clinicians to provide patients with an accurate prognostic outlook, or estimated timeline of disease progression. As such, recent efforts have been made to improve prognostic prediction models. Long and colleagues developed the Prognostic Index Normed (PIN) score, which incorporates CAG repeat length, age, as well as two clinical measures, the Unified Huntington's Disease Rating Scale (UHDRS) Total Motor Score (TMS), and the Symbol Digit Modalities Test (SDMT) score.[5] In addition, we have shown that plasma levels of neurofilament light (NfL), a neuronal marker associated with neurodegeneration, is associated with predicted years to CMD, as determined using a participant's CAG repeat length and age,[2,6] as well as when determined using the PIN score.[5,7] We have proposed that a plasma NfL cut-point of ≥45.0 pg/ml could distinguish participants who are within ten years until CMD.[6,7] While these efforts have contributed to predicting years to CMD, the division of patients into 'premanifest' and 'manifest' HD categories has itself inappropriately simplified dynamic sub-clinical and clinical disease progression into two broadly defined subgroups. The limitations of using these subgroups has been acknowledged by the HD research community, and recently a new biological classification system—the

Huntington's Disease Integrated Staging System (HD-ISS)—has been proposed.[8] The aim of the HD-ISS, developed by Tabrizi et al.,[8] is to incorporate HD pathophysiology and biomarker changes that precede overt functional and motor symptom presentation and a clinical diagnosis of manifest HD. Briefly, the HD-ISS is comprised of four staging groups, with participants in Stage 0 defined as those carrying the *HTT* mutation, but without symptom presentation or detectable pathological change; those in Stage 1 exhibiting underlying basal ganglia pathology as measured by magnetic resonance imaging; those in Stage 2 displaying a clinical phenotype as measured by changes on the SDMT and TMS; and those in Stage 3 demonstrating functional decline as measured by changes on the UHDRS Total Functional Capacity (TFC) score and Independence score.[8] To be considered as a potential landmark for HD-ISS categorisation, a measure must have been studied longitudinally, in a cohort of 100 or more participants, and must be predictive of disease progression, as shown by two independently generated models or a single model that used two independent cohorts.[8] To date, no wet biomarkers have met this stringent criteria.[8] In addition, many observational research studies lack the ability to collect quantitative neuroimaging data; while this does not preclude these sites from screening for clinical trial inclusion, with trial eligibility criteria predominantly requiring Stage 2 or Stage 3-equivalent membership, a lack of quantitative neuroimaging data would prevent research sites from utilising the HD-ISS for non-clinical trial studies of early disease progression. Consequently, the authors have additionally provided PIN value thresholds that would maximise separation amongst the stages, which we will use as approximate proxies for

HD-ISS classification.[9] The aim of the current study was to assess the cross-sectional differences between select clinical measures, as well as plasma NfL levels, and the PIN score-approximated HD-ISS stages, in a unique, well-characterised patient population from a single study site, to determine whether consideration of plasma NfL levels might be useful for enriching stages of the HD-ISS.

Methods

Ethics
This study was approved by the University of California, San Diego (UCSD) Institutional Review Board (IRB) Committee (IRB Protocol #170038), in accordance with the requirements of the Code of Federal Regulations on the Protection of Human Subjects. All participants gave written informed consent prior to sample collection.

Human subjects
Carriers of the *HTT* gene mutation with a family history of the disorder (herein referred to as gene mutation carriers), as well as healthy controls (HC) who were not at risk of inheriting the *HTT* mutation, were recruited from the UCSD Huntington's Disease Society of America (HDSA) Center of Excellence (CoE). Gene mutation carriers were excluded if they were participating in a clinical trial for HD. All HC had a Montreal Cognitive Assessment (MoCA) score of 26 or higher. No other inclusion or exclusion criteria were applied. All individuals who were eligible, had consented, and attended a clinical assessment to provide clinical data and blood samples at UCSD HDSA CoE, were included. Gene mutation carriers were categorised into HD-ISS stages, as outlined below. Demographic and disease data were collected at the time of sample collection, including gender, age, CAG repeat length, years of education, and family history. Gender was self-reported by the participant.

Clinical assessment
Study participants underwent a clinical assessment at the time of their study visit, which included cognitive testing and functional measures, and motor ratings. The cognitive battery included the Mini-Mental State Examination (MMSE; score range 0–30),[10] MoCA (score range 0–30),[11] SDMT (score range 0–110)[12] and Stroop Word Reading test (SWR). Functional capacity was evaluated using the UHDRS TFC (score range 0–13), and Independence (score range 0–100) ratings.[13] Motor dysfunction was assessed using the UHDRS TMS (score range 1–124).[13] The SWR, SDMT, TFC and TMS were also incorporated into the composite UHDRS score (cUHDRS) as an additional measure of disease burden.[14] PIN scores were calculated using a previously published formula, which incorporates a participant's age at the time of appointment, CAG repeat length, SDMT score, and TMS score.[5]

HD-ISS categorization
Participant HD-ISS membership was approximated according to their PIN score. PIN score thresholds have been provided as one example of rough proxies when the quantitative neuroimaging variables required for the HD-ISS are not available.[8,9,15] Specifically, participants with a PIN score ≤ -0.34 were categorised as Stage 0, those with a PIN score of > -0.34 to 0.60 as Stage 1, those with a PIN score of >0.60–2.31 as Stage 2 and those with a PIN score greater than 2.31 as Stage 3. Three participants refused to complete the SDMT at their appointment and were therefore unable to be grouped by PIN score. As CAG-Age Product (CAP; calculated as age × (CAG repeat length − 33.66)) score proxies for the HD-ISS stage thresholds have also been estimated,[9] these participants were staged based on their CAP score instead. In order to compare the accuracy of PIN-substituted HD-ISS categorisation, to categorisation using the originally proposed HD-ISS,[8] participants were also staged using this latter method. As quantative neuroimaging data was not available, HD-ISS stages 0 and 1 were combined for this latter comparison.

Plasma collection and NfL analysis
Blood was drawn by venipuncture into 2 ml lavender/EDTA tubes. EDTA/whole blood was mixed by inversion and centrifuged (900*g*, 15 min). The supernatant was isolated, aliquoted into 1 ml aliquots, snap frozen and stored at −80 °C. Plasma levels of NfL were measured in duplicate using a Meso Scale Discovery (MSD; Rockville, MD) R-Plex Assay (Cat# F217X; researcher determined lower limit of detection: 2.3 pg/ml) as previously described.[6]

Statistics
Analyses were conducted with GraphPad Prism version 8.4.2 for Windows (GraphPad Software, La Jolla, CA, USA). Due to a non-normal distribution of data, cohort characteristics were compared using Kruskal–Wallis and Mann–Whitney U tests. As participants were placed in HD-ISS stages based on their PIN score, and the PIN score equation accounts for age and CAG repeat length, additional adjustments for these covariates were not made.[5,8,9,15]

Role of funders
The funding sources had no role in study design, conduct, analysis, interpretation or writing of the manuscript, or in the decision to submit the manuscript.

Results

Participant characteristics
A total of 290 plasma samples were collected from 181 participants: 50 Stage 0, 64 Stage 1, 63 Stage 2, 63 Stage 3, and 50 HC samples. Cohort demographic data is

summarised in Table 1. Cohorts differed significantly by age; the Stage 0 cohort was significantly younger than all other cohorts (Dunn's post-hoc $p < 0.0001$ for all comparisons), and Stage 1 was significantly younger than Stage 2 and Stage 3 (Dunn's post-hoc $p = 0.009$ and $p = 0.02$, respectively). The Stage 3 cohort differed significantly in CAG repeat length compared to Stage 0 (Dunn's post-hoc $p < 0.0001$), Stage 1 (Dunn's post-hoc $p = 0.003$) and Stage 2 (Dunn's post-hoc $p = 0.02$); Stage 0 also differed significantly in CAG repeat length when compared to Stage 2 (Dunn's post-hoc $p = 0.04$).

Considering cognition, as measured by the MoCA, significant differences were present between cohorts (Table 1). Post-hoc analyses showed that all analysed cohorts were significantly different from each other, except for the HC cohort compared to Stage 0, and Stage 1 compared to Stage 2 (Supplementary Fig. S1A). Similarly, significant differences were present between cohorts when cognition was measured using the MMSE (Table 1). Post-hoc analyses showed that all analysed cohorts were significantly different from each other, except for Stage 0 compared to HC and compared to Stage 1, and Stage 1 compared to Stage 2 (Supplementary Fig. S1B). Composite UHDRS scores, commonly used as a clinical trial outcome measure, were also significantly different across cohorts (Table 1). Post-hoc analyses showed that all analysed cohorts were significantly different from each other, except for HC compared to Stage 0 and compared to Stage 1 (Supplementary Fig. S1C). Cohort demographic data from unique participants is outlined in Supplementary Table S1.

HD-ISS group membership by PIN

HD-ISS membership into each group assumes that conditions for membership into all lower groups have been met.[8,9] For example, membership into Stage 3 assumes that all included participants present with cognitive and motor symptoms, as measured by the SDMT and TMS, above the threshold for Stage 2. In addition, functional symptoms, as measured by the UHDRS TFC and Independence score, would need to meet the threshold for Stage 3 membership (<13 for TFC and <100 for Independence, respectively). We thus examined TMS, SDMT, TFC and Independence scores by stage, to characterise the accuracy of membership in our PIN score-approximated cohorts. All participants in Stage 3 had SDMT (Fig. 1a) and TMS scores (Fig. 1b) within or above the threshold for Stage 2 membership; however not all Stage 2 participants met the SDMT or TMS scores for Stage 2 membership. SDMT scores differed significantly between stages (Kruskal–Wallis $\chi 2 = 175.1$, $p < 0.0001$), with significant post-hoc variability detected between all cohorts (Fig. 1a). TMS scores differed significantly between all stages (Kruskal–Wallis $\chi^2 = 189.5$, $p < 0.0001$), with significant post-hoc variability similarly detected between all cohorts (Fig. 1b). Regarding functional change, one Stage 3 participant had a TFC score of 13, with all others below 13 (Fig. 1c), and five Stage 3 participants had Independence scores of 100, with all others below 100 (Fig. 1d). Thresholds proposed for the original HD-ISS were considered and presented as ranges (Fig. 1a–b), rather than fixed values, as the HD-ISS staging thresholds, when used without PIN score approximation, vary by age and education level.[8]

Predictors of years to CMD by HD-ISS categorisation

We have previously shown that predicted years to CMD at 60% probability (yCMD), as determined by either Langbehn and colleagues' formula[2] or PIN-derived,[5] are significantly correlated with plasma NfL levels.[6,7] As the PIN-derived yCMD measure is highly correlated with Langbehn and colleagues' measure, both relate to plasma NfL levels,[2] and the PIN score is inherent to the PIN score-approximated staging system, we limited our comparisons to plasma NfL and Langbehn and colleagues' formula.

Plasma NfL levels differed significantly by HD-ISS stage (Kruskal–Wallis $\chi^2 = 120.2$, $p < 0.0001$); post-hoc

	HC	S0	S1	S2	S3	χ^2, p
n	50	50	64	63	63	
Age [years]	57.5, 37.8–65.0	35.0, 30.5–40.0	45.5, 37.3–58.0	56.0, 49.0–65.0	56.0, 47.0–65.0	80.3, <0.0001
Gender n[F/M]	22/28	27/23	31/33	24/39	34/29	4.45, 0.35
Education [years]	16.0, 13.8–16.3	16.0, 13.5–16.5	16.0, 14.0–17.0	16.0, 13.0–18.0	14.0, 12.0–17.0	7.7, 0.10
CAG	NA	40.0, 40.0–42.0	42.0, 40.0–43.8	42.0, 41.0–43.0	43.0, 41.0–46.0	31.2, <0.0001
DCL	NA	0.0, 0.0–0.0	1.0, 0.0–3.0	4.0, 2.0–4.0	4.0, 4.0–4.0	145.6, <0.0001
MoCA	28.0, 27.0–29.0	28.0, 27.0–29.0	27.0, 26.0–28.0	26.0, 24.0–28.0	22.0, 18.0–26.0	90.5, <0.0001
MMSE	29.0, 28.0–30.0	29.0, 28.0–29.3	28.0, 26.0–29.0	28.0, 26.0–29.0	24.0, 22.0–27.0	81.0, <0.0001
cUHDRS	17.3, 16.5–18.0	18.0, 17.1–18.8	16.3, 15.8–17.3	14.4, 13.6–16.1	9.2, 6.9–10.7	201.1, <0.0001

HC, healthy controls; S0–S3, Stage 0–Stage 3; DCL, Diagnostic Confidence Level; MoCA, Montreal Cognitive Assessment; MMSE, Mini-Mental State Examination; cUHDRS, composite Unified Huntington's Disease Rating Scale; NA, not applicable. All continuous analyses conducted using Kruskal–Wallis test; comparison of gender proportions conducted using a Chi Square test.

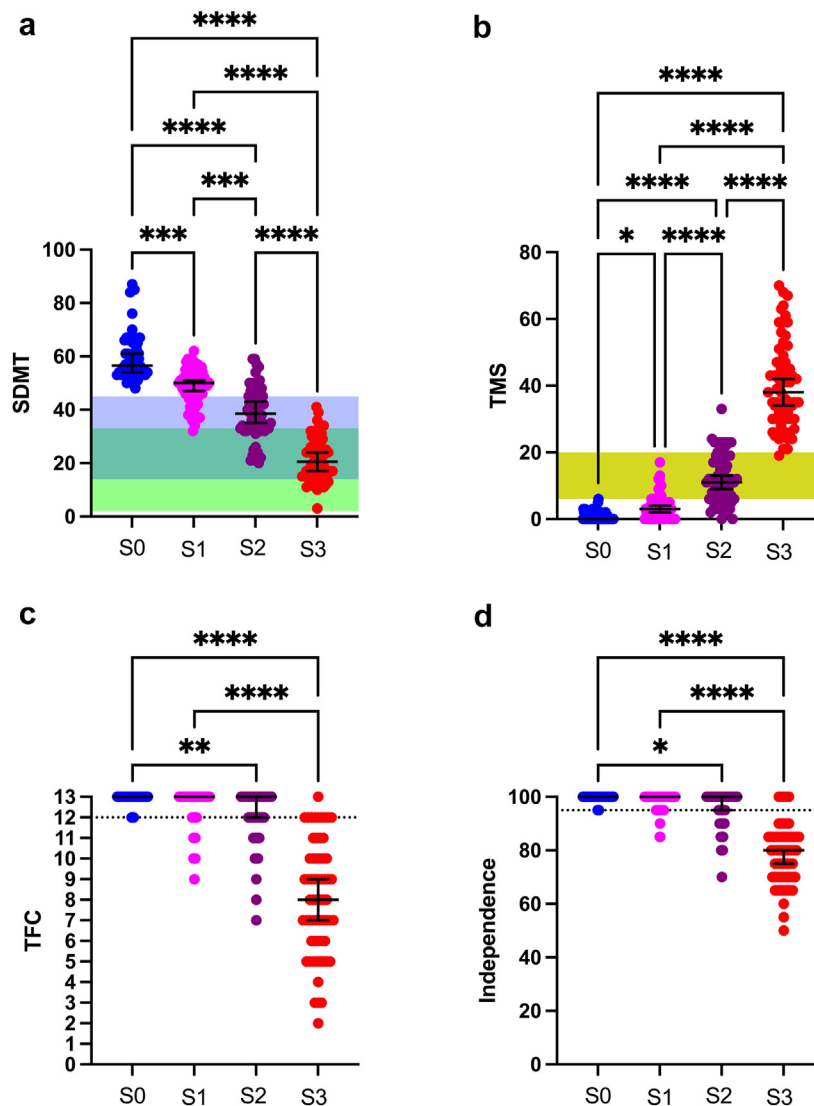*Table 1:* Cohort demographic comparisons [median, interquartile range].

Fig. 1: A comparison of measures incorporated into the HD-ISS—the Symbol Digit Modalities Test (SDMT, panel a), Total Motor Score (TMS, panel b), Total Functional Capacity (TFC, panel c) and Independence score (panel d), grouped by participant membership into the PIN score-approximated HD-ISS. Colored bars (a, b) indicate the proposed threshold range for each measure, which vary by age and, for SDMT, education level (Panel a: green = less than or equal to high school education; blue = greater than high school education). TFC scores are required to be <13, and Independence scores are required to be <100, for membership into Stage 3 (dotted line in c, d), with lower cut-offs required for older age groups (not shown). Comparisons conducted using Kruskal–Wallis test with Dunn's post-hoc comparisons. Error bars represent median and interquartile range; ****$p < 0.0001$, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. Sample sizes: Stage 0 (S0) = 50, Stage 1 (S1) = 64, Stage 2 (S2) = 63, Stage 3 (S3) = 64, Healthy Controls (HC) = 50.

analyses indicated that Stage 3 participants had significantly higher levels compared to HC, Stage 0, Stage 1 (all Dunn's post-hoc ps < 0.0001) and Stage 2 ($p = 0.0002$). In addition, Stage 2 participants had significantly higher levels compared to HC, Stage 0 (both Dunn's post-hoc ps < 0.0001), and Stage 1 (Dunn's post-hoc $p = 0.008$), and Stage 1 participants had significantly higher levels compared to Stage 0 (Dunn's post-hoc $p = 0.03$) (Fig. 2a). Cohort medians were above our previously proposed

plasma NfL cut-point of $\geq$45.0 pg/ml for $\leq$10 yCMD, for Stage 1 (median of 48.7 pg/ml), Stage 2 (median of 67.0 pg/ml), and Stage 3 (median of 117.6 pg/ml), but below this cut-point for HC (median of 27.3 pg/ml) and Stage 0 (median of 22.1 pg/ml) (Fig. 2a). yCMD also differed significantly across cohorts (Kruskal–Wallis $\chi^2 = 185.4$, $p < 0.0001$); post-hoc analyses showed that Stage 0 participants had significantly more yCMD compared to Stage 1, Stage 2, and Stage 3 (all post-hoc
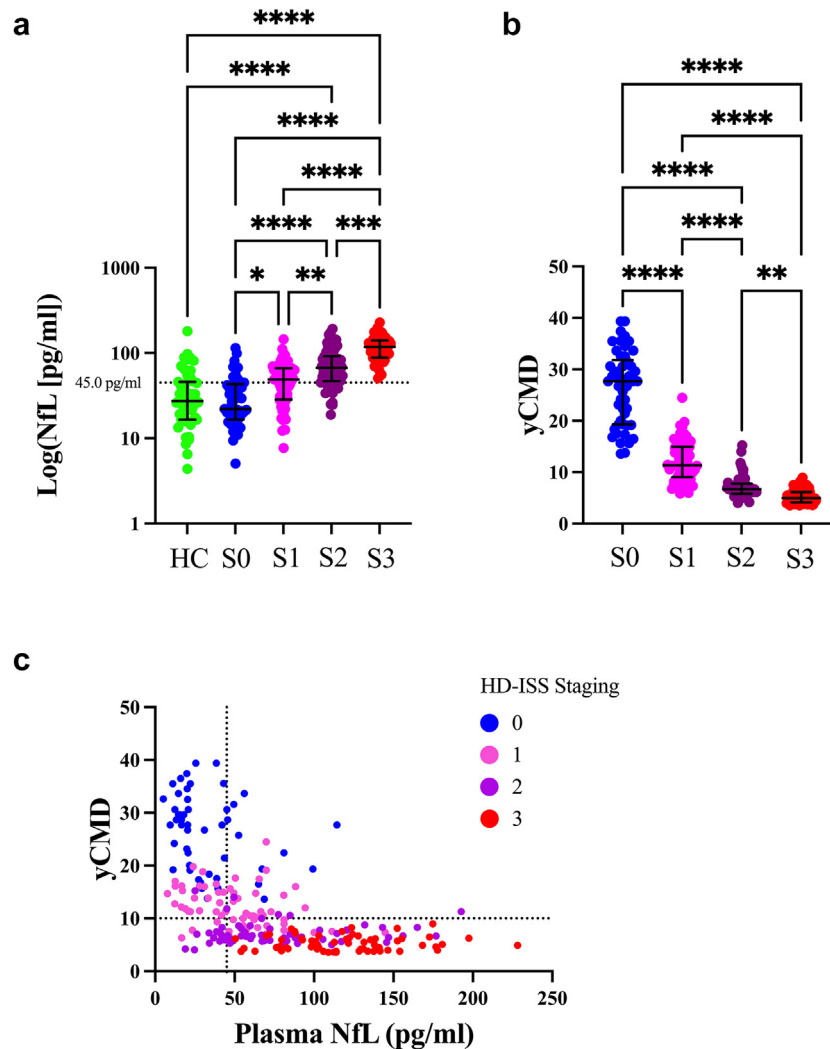
Fig. 2: Plasma NfL levels (a) and predicted years to 60% probability of clinical motor diagnosis (yCMD) (b) differed between HD-ISS stages and healthy controls (HC). Comparisons conducted using Kruskal–Wallis test with Dunn's post-hoc comparisons. Error bars represent median and interquartile range; ****$p < 0.0001$, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. (c) Depicts the relationship between yCMD and plasma NfL levels, including our proposed cut-point of 45.0 pg/ml (dotted lined in x-axis) for identifying participants with less than ten yCMD (dotted line in y-axis), and the HD-ISS. Sample sizes: Stage 0 (S0) = 50, Stage 1 (S1) = 64, Stage 2 (S2) = 63, Stage 3 (S3) = 64, Healthy Controls (HC) = 50. All measurements conducted with two biological replicates.

ps < 0.0001); Stage 1 participants also had significantly more yCMD compared to Stage 2 and Stage 3 (all Dunn's post-hoc ps < 0.0001), and Stage 2 had significantly more yCMD compared to Stage 3 ($p = 0.004$) (Fig. 2b). The relationship between yCMD, plasma NfL levels, and HD-ISS is also depicted in Fig. 2c.

Comparing original HD-ISS versus PIN score substituted HD-ISS categorisation

In order to determine the accuracy of PIN score substitution for HD-ISS categorisation, particularly as it relates to plasma NfL levels, we compared stage membership using both the original HD-ISS formula, as well as with PIN score substitution. Membership into HD-ISS Stage 1 is governed by quantitative neuroimaging landmarks, and therefore for the purpose of this analysis, Stage 0 and Stage 1 were combined. A chi square comparison of membership showed significant differences between expected (the original HD-ISS) versus observed (PIN-substituted HD-ISS) values ($\chi^2 = 19.4$, $p < 0.0001$). More participants were assigned to PIN-substituted HD-ISS Stage 2, instead of PIN-substituted HD-ISS Stage 3, than expected (Table 2). A breakdown of membership by original HD-ISS, compared to PIN-substituted HD-ISS categorisation, is also presented in Table 3.

	Expected (original HD-ISS) (% total)	Observed (PIN-substituted HD-ISS) (% total)
Stage ≤1	109 (45.4)	114 (47.5)
Stage 2	31 (12.9)	63 (26.3)
Stage 3	100 (41.7)	63 (26.3)

PIN, Prognostic Index Normed; HD-ISS, Huntington's Disease Integrated Staging System.

Table 2: **Comparison of participant expected (original HD-ISS) versus observed (PIN-substituted HD-ISS) membership.**

A within-stage comparison of plasma NfL levels by PIN-substituted, and original HD-ISS, categories showed significant differences between Stage 3 values (Mann–Whitney U = 2458, p = 0.02). No other stages displayed significant differences ([Fig. 3](#)).

Discussion

We have previously reported associations between PIN scores, PIN-derived yCMD, CAG- and age-derived yCMD, and plasma NfL levels. Importantly, we have proposed that a plasma NfL cut-point of ≥45.0 pg/ml distinguishes patients who are within ten yCMD.[6,7] In the current study, we showed that HD-ISS stages differed significantly by plasma NfL levels and yCMD. Notably, Stage 1 participants displayed median plasma NfL levels just above our previously proposed cut-point for being ≤yCMD (45.0 pg/ml)[6,7] at 48.70 pg/ml, Stage 2 and 3 were above this cut-point, whereas HC and Stage 0 were below this cut-point. The detection of plasma NfL levels above our proposed cut-point in approximately 50% of Stage 1 participants is interesting, as this suggests biochemical change in the absence of clinical symptom presentation. While this is not surprising, given that membership into Stage 1 was first designated by the presence of asymptomatic neuroimaging change,[8] and plasma NfL levels have previously been associated with caudate and putamen volume,[16,17] our findings would suggest that plasma NfL—and potentially other biofluid biomarkers—may be useful for enriching HD-ISS categorisation. The association between plasma NfL levels and HD-ISS stages is also reflected in our comparison of yCMD and HD-ISS staging; Stage 0 participants had a median of 27.7

	Original HD-ISS category		
	Stage ≤1	Stage 2	Stage 3
PIN-substituted HD-ISS category			
Stage ≤1	97	7	10
Stage 2	12	22	29
Stage 3	0	2	61

PIN, Prognostic Index Normed; HD-ISS, Huntington's Disease Integrated Staging System.

Table 3: **Participant membership by HD-ISS categorization method.**
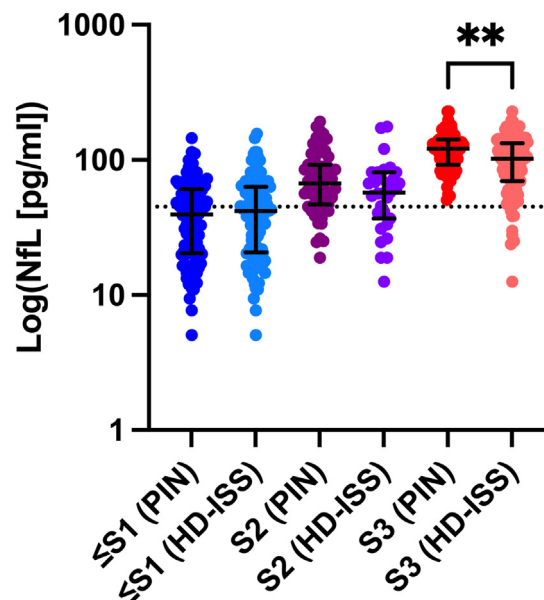

Fig. 3: Plasma NfL levels differed significantly by HD-ISS categorisation method (original HD-ISS thresholds [HD-ISS] versus PIN-substituted thresholds [PIN]) for Stage 3 (S3), but not Stage 2 (S2) or Stage 0 and Stage 1 combined (≤S1). Error bars represent median and interquartile range. The dotted line indicates our proposed plasma NfL cut-point (45.0 pg/ml) for <10 predicted years to clinical motor diagnosis. Sample sizes: ≤Stage 1 (≤S1) (PIN) = 114, ≤S1 (HD-ISS) = 109; Stage 2 (S2) (PIN) = 63; S2 (HD-ISS) = 31; Stage 3 (S3) (PIN) = 67, S3 (HD-ISS) = 100. All measurements conducted with two biological replicates.

yCMD, Stage 1 had a median of 11.3 yCMD, and Stage 2 and 3 had less than ten yCMD with 60% probability of onset (6.7 and 5.0 yCMD, respectively). Further, a preliminary post-hoc analysis of Stage 1 participants who were below versus above our plasma NfL cut-point of 45.0 pg/ml, revealed that these groups differed significantly in age (median of 40.0 vs 50.5, respectively, p = 0.03), PIN score (median of −0.05 vs 0.20, respectively, p = 0.007), and yCMD (median of 13.4 vs 10.4 years, respectively, p = 0.01).

It has previously been reported that participants with a PIN score of zero at baseline have a yCMD of ten or less years at a probability of 50%,[8,18] with PIN scores less than zero indicating more than ten years until CMD. When considering the PIN score-approximated HD-ISS stages, Stage 0 includes participants with PIN scores of ≤−0.34, and Stage 1 includes participants with PIN scores of >−0.34 to 0.60. This again shows that Stage 1 contains a mix of participants above and below ten years to CMD. These findings suggest that consideration of plasma NfL levels would therefore assist in further characterisation of Stage 1 participants for prognostic and research purposes.

Another limitation of using the PIN score alone for approximation of HD-ISS staging, is the HD-ISS

assumption that if any stage conditions in later stages are met (for example, SDMT/cognitive and TMS/motor symptoms for membership into Stage 2), that the Stage 1 conditions (for example, putamen and/or caudate volume below the specified threshold) are also met. However, a participant's PIN score is determined by their age, CAG repeat length, SDMT and TMS score through (PIN = 51 × TMS − 34 × SDMT + 7 × Age × (CAG − 34) − 883) + 1044)) and consequently, this PIN score would increase as their lifetime exposure to the *HTT* mutation, a function of their age and CAG repeat length, increases. As such, it is possible that by using PIN scores alone, a number of participants will be placed, for example in Stage 2, while still asymptomatic. While the age and CAG repeat length of these participants may place them at higher risk of developing symptoms in the near future, such as during the span of a clinical trial, they remain exceptions to the intended HD-ISS staging criteria. It is important to note at this point that some overlap is expected between stages, which has been acknowledged by Tabrizi and colleagues with publication of the HD-ISS,[8,9,15] due to the variability in disease progression that is not accounted for by HD-ISS or PIN-included measures, and therefore stage membership does not always match a participant's global clinical status. For example, when using the PIN score, or another substitute, the top percentile of Stage 2 TMS values will overlap with the bottom percentile of Stage 3 TMS values.[8,9,15] Overall, the purpose of the HD-ISS is to enrich selection criteria for the conduct of clinical trials, and other research studies, and not provide 100% specificity.

In order to assess the accuracy of PIN-substituted HD-ISS categorisation, we compared participant membership to that determined using the originally proposed HD-ISS, with Stage 0 and 1 combined (termed '≤Stage 1') to circumvent the absence of quantitative neuroimaging data.[8] Our findings show that these two categorisation methods only diverge when it comes to membership in Stage 2, versus Stage 3, perhaps since Stage 3 membership in the original HD-ISS is also governed by functional (UHDRS TFC and Independence score) decline. The PIN-substituted HD-ISS, on the other hand, is governed by SDMT, TMS, age and CAG repeat, and therefore does not consider functional decline. Importantly, PIN-substituted, and original HD-ISS, membership did not differ for the ≤Stage 1 cohorts, and both original and PIN-substituted groups included a number of participants with plasma NfL levels above our proposed cut-point for ≤10 (45.0 pg/ml) yCMD.

Our analysis of plasma NfL levels, and yCMD, staged by PIN-derived HD-ISS, suggests that while there may not be complete consistency between group membership when assigned using the original clinical measure-derived HD-ISS, versus the PIN score-approximated HD-ISS, those staged as near or at

CMD for the purpose of clinical trial recruitment—those in Stage 2 and 3—would be recruited regardless of HD-ISS staging method. Indeed, a suggested alternative approach for utilisation of the original HD-ISS in the absence of quantitative neuroimaging data, is to stage participants into Stage <2, 2, or 3, particularly as all current clinical trials focus on HD-ISS Stage 2 or 3.[8,9,15] While our study was not powered or intended to support the use of plasma NfL, or inclusion of Stage 1 participants, in clinical trial selection criteria, we contend that plasma NfL may be useful in identifying Stage 1 participants close to CMD for research purposes.

It is also important to note that use of the PIN score to estimate HD-ISS categorisation is just one exemplified method provided by the authors of the HD-ISS, and while we found good concordance in our cohort when compared to use of the original HD-ISS, this may not be the case for all studies or sites. The authors of the HD-ISS presented the PIN score thresholds as one example, however additionally focused on a machine-learning imputation method.[9] This imputation method has been incorporated into the latest Enroll-HD dataset periodic data set (PDS6). While we were not able to apply this dataset for the current study, we encourage researchers to utilise it when possible.

There are a number of considerations and limitations associated with this investigation. We used data obtained from the same research site and using the same methodology as our previous NfL studies[6,7] and therefore, analytical and clinical validation of these findings in a unique cohort of individuals with HD, using the same, as well as alternative NfL assays, would be extremely valuable. In addition, the Meso Scale Discovery assay utilised in our study is part of the manufacturer's R-Plex platform; this means the assay has not been validated by the manufacturer, nor standardised to NfL assays on other manufacturer platforms. Moreover, we have previously reported up to 22% inter-assay variation in plasma NfL values for samples included in this study, after normalisation across assay plates,[6] and others have reported discrepancies in absolute quantification values between assays from different manufacturers.[19] Other factors may also affect plasma NfL levels such as kidney disease and body composition (e.g lower plasma NfL levels in participants with higher BMI).[20] Therefore, more work needs to be done, both by researchers and manufacturers, to improve, validate and standardise analyses of biofluid markers such as NfL, if they are to be used as prognostic biomarkers. Until such validation has occurred, our plasma NfL cut-point for ≤10 yCMD cannot be considered as the norm for the field.

It is also important to reiterate that the PIN score thresholds used to approximate HD-ISS membership are proxies, which have been selected to maximise the separation between stages.[9] These thresholds are not intended, and have not been validated, as direct

substitutes for the collection and use of imaging data. We acknowledge that our findings would be strengthened by an analysis of plasma NfL levels and quantitative neuroimaging markers, categorised by PIN-substituted HD-ISS. Quantitative neuroimaging of such a population, with a sample size large enough to avoid a type II error, is not feasible for many research sites, however we do believe this is important research that needs to be conducted. Overall, while our findings are exploratory and require external optimisation and validation, including confirmation of our prognostic cut-point in a longitudinal study, we contend that plasma NfL levels, may be useful for the enrichment of Stage 1 participants who are closer to CMD.

Our findings suggest that the HD-ISS is an appropriate means to stratify participants based on symptom presentation, and for research purposes. However, the additional use of biomarkers alongside the HD-ISS may help enrich or define subgroups within each stage. We propose that additional longitudinal studies of biofluid markers, including plasma NfL, with sufficient sample sizes, be conducted to fully assess the potential use of these markers alongside the HD-ISS. Promisingly a number of longitudinal biomarker studies are currently underway.

Contributors
Conceptualization, E.A.T. and J.C-B.; Formal analysis, G.M.P.; Funding acquisition, E.A.T. and J.C-B.; Supervision, E.A.T. and J.C-B.; Writing—original draft, G.M.P.; Writing—review & editing, E.A.T. and J.C-B. All authors read and approved the final version of the manuscript. G.M.P., E.A.T. and J.C-B. have verified the underlying data.

Data sharing statement
Anonymised summary data are available from the corresponding author by reasonable formal request from qualified researchers, subject to a data sharing agreement and in compliance with the requirements of the funding bodies and institutions.

Declaration of interests
J.C-B. has participated in Speakers Bureaus for Teva Pharmaceuticals and EMD Serono; J.C-B. has participated on a Data Safety Monitoring Board for UniQure. The authors have no other conflicts of interests to disclose.

Appendix A. Supplementary data
Supplementary data related to this article can be found at https://doi.org/10.1016/j.ebiom.2023.104646.

References
1. Andresen JM, Gayán J, Djoussé L, et al. The relationship between CAG repeat length and age of onset differs for Huntington's disease patients with juvenile onset or adult onset. *Ann Hum Genet*. 2007;71(3):295–301.
2. Langbehn DR, Hayden MR, Paulsen JS, PREDICT-HD Investigators. 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;153(2):397–408.
3. Wexler NS, Lorimer J, Porter J, et al. Venezuelan kindreds reveal that genetic and environmental factors modulate Huntington's disease age of onset. *Proc Natl Acad Sci U S A*. 2004;101(10):3498–3503.
4. Arrasate M, Finkbeiner S. Protein aggregates in Huntington's disease. *Exp Neurol*. 2012;238(1):1–11.
5. Long JD, Langbehn DR, Tabrizi SJ, et al. Validation of a prognostic index for Huntington's disease. *Mov Disord*. 2017;32(2):256–263.
6. Parkin GM, Corey-Bloom J, Snell C, Castleton J, Thomas EA. Plasma neurofilament light in Huntington's disease: a marker for disease onset, but not symptom progression. *Parkinsonism Relat Disord*. 2021;87:32–38.
7. Parkin GM, Corey-Bloom J, Long JD, Snell C, Smith H, Thomas EA. Associations between prognostic index scores and plasma neurofilament light in Huntington's disease. *Parkinsonism Relat Disord*. 2022;97:25–28.
8. 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.
9. Long JD, Gantman EC, Mills JA, et al. Applying the Huntington's disease integrated staging system (HD-ISS) to observational studies. *J Huntingtons Dis*. 2023;12(1):57–69.
10. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189–198.
11. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53(4):695–699.
12. Smith A. Symbol digit modalities test. In: Spreen O, Strauss E, eds. *A compendium of neuropsychological tests*. 2nd ed. Los Angeles: Western Psychological Services; 1982.
13. Kieburtz K, Penney JB, Corno P, et al. Unified Huntington's disease rating scale: reliability and consistency. *Neurology*. 2001;11(2):136–142.
14. 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.
15. Mills JA, Long JD, Vaidya J, Sampaio C, Sathe S. *Comorbidies in Huntington's disease: an enroll-HD analysis*. International Congress of Parkinson's Disease and Movement Disorders; 2022. Madrid, Spain.
16. 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.
17. Byrne LM, Rodrigues FB, Johnson EB, et al. Evaluation of mutant huntingtin and neurofilament proteins as potential markers in Huntington's disease. *Sci Transl Med*. 2018;10(458):eaat7108.
18. Langbehn DR, Hersch S. Clinical outcomes and selection criteria for prodromal huntington's disease trials. *Mov Disord*. 2020;35(12):2193–2200.
19. Kuhle J, Barro C, Andreasson U, et al. Comparison of three analytical platforms for quantification of the neurofilament light chain in blood samples: ELISA, electrochemiluminescence immunoassay and Simoa. *Clin Chem Lab Med*. 2016;54(10):1655–1661.
20. Arslan B, Zetterberg H. Neurofilament light chain as neuronal injury marker–what is needed to facilitate implementation in clinical laboratory practice? *Clin Chem Lab Med*. 2023. https://doi.org/10.1515/cclm-2023-0036.