



OPEN

Quantifying mutant huntingtin protein in human cerebrospinal fluid to support the development of huntingtin-lowering therapies

Stephanie Vauleon¹, Katharina Schutz¹, Benoit Massonnet¹, Nanda Gruben², Marianne Manchester³, Alessandra Buehler¹, Eginhard Schick¹, Lauren Boak⁴ & David J. Hawellek⁵✉

Huntington's disease (HD) is caused by a cytosine adenine guanine-repeat expansion in the huntingtin gene. This results in the production of toxic mutant huntingtin protein (mHTT), which has an elongated polyglutamine (polyQ) stretch near the protein's N-terminal end. The pharmacological lowering of mHTT expression in the brain targets the underlying driver of HD and is one of the principal therapeutic strategies being pursued to slow or stop disease progression. This report describes the characterisation and validation of an assay designed to quantify mHTT in the cerebrospinal fluid of individuals with HD, for use in registrational clinical trials. The assay was optimised, and its performance was characterised with recombinant huntingtin protein (HTT) varying in overall and polyQ-repeat length. The assay was successfully validated by two independent laboratories in regulated bioanalytical environments and showed a steep signal increase as the polyQ stretch of recombinant HTTs pivoted from wild-type to mutant protein forms. Linear mixed effects modelling confirmed highly parallel concentration–response curves for HTTs, with only a minor impact of individual slopes of the concentration–response for different HTTs (typically < 5% of the overall slope). This implies an equivalent quantitative signal behaviour for HTTs with differing polyQ-repeat lengths. The reported method may be a reliable biomarker tool with relevance across the spectrum of HD mutations, which can facilitate the clinical development of HTT-lowering therapies in HD.

Huntington's disease (HD) is a rare, genetic neurodegenerative disease that is characterised by a triad of cognitive, behavioural and motor symptoms^{1,2}. Initial changes in brain pathophysiology underlie the early stage of HD. As the disease progresses, cognitive and motor symptoms become clinically detectable, followed by continued decline in body function^{3,4}.

HD is caused by a cytosine adenine guanine (CAG)-repeat expansion in the huntingtin gene (*HTT*), which is a direct determinant of potential or confirmed HD onset. A CAG-repeat length of ≥ 40 causes HD, while a CAG-repeat length of ≤ 26 does not. The middle ranges of 27–35 and 36–39 CAG repeats are known as 'intermediate' and 'reduced penetrance' respectively; the former will not cause HD but is associated with an increased risk of HD in subsequent generations, and the latter may or may not cause HD in the individual's lifetime^{1,5–7}.

The CAG-repeat expansion results in the production of toxic mutant huntingtin protein (mHTT), which has an elongated polyglutamine (polyQ) stretch near the protein's N-terminal end^{1,5,8}. Levels of mHTT in cerebrospinal fluid (CSF) correlate with disease stage, symptom severity and markers of neuronal damage in people with HD^{9,10}. Lowering mHTT production, via the degradation of *HTT* mRNA for example, targets the underlying driver of HD and interferes with the direct causal pathway of the disease. Consequently, mHTT is a key biomarker of HD as it has a direct causal involvement in the pathophysiology of the disease. This renders it a direct target for pharmacological interventions⁵.

¹Regulated Bioanalysis, Pharmaceutical Sciences, Pharma Research and Early Development (pRED), F. Hoffmann-La Roche Ltd, Basel, Switzerland. ²ICON Bioanalytical Laboratories, Assen, The Netherlands. ³Pharmaceutical Sciences, Pharma Research and Early Development (pRED), F. Hoffmann-La Roche Ltd, Basel, Switzerland. ⁴F. Hoffmann-La Roche Ltd, Basel, Switzerland. ⁵Biomarkers and Translational Technologies, Pharma Research and Early Development (pRED), F. Hoffmann-La Roche Ltd, Basel, Switzerland. ✉email: david.hawellek@roche.com

Huntingtin protein (HTT) has an extremely low abundance in the CSF¹¹, making ultra-sensitive platforms the most suitable method for detection. A novel, ultra-sensitive single molecule counting (SMC) mHTT immunoassay on the Erenna[®] platform was shown by Wild et al.⁹ to quantify CSF mHTT in association with proximity to disease onset and reductions in cognitive and motor function. Additionally, a novel, ultra-sensitive mHTT detection assay was developed by Southwell et al.¹¹, adapted from the highly sensitive protein detection technique of microbead-based immunoprecipitation followed by flow cytometry (IP-FCM). This HTT IP-FCM assay accurately detected mHTT in the CSF of Hu97/18 mice and individuals with HD, whereby CSF mHTT levels increased with disease stage and decreased after HTT suppression.

Initial measurements of clinical samples with the SMC assay in a research-grade environment suggest it may support the application of mHTT quantification as a biomarker in HD clinical trials for HTT-lowering therapies^{12,13}. A research-grade version of the assay has been used to analyse HD samples from the Phase I/IIa study of the antisense-oligonucleotide (ASO) tominersen (NCT02519036)¹³.

The ligand binding assay uses capture antibody 2B7 which binds to both mHTT and wild-type HTT (wtHTT), and detection antibody MW1 which binds to the extended polyQ stretch of mHTT (Fig. 1). Although MW1 is used in many available assays, the specificity of MW1 for mHTT remains relatively unclear. MW1 may have differential binding properties depending on the sub-cellular location of mHTT¹⁴ as well as the exact number of CAG repeats in HTT^{11,12,15,16} suggesting that not all HTT species are equally detected by MW1. The current study extensively characterised the assay response via experiments with recombinant HTT and patient CSF across a wide range of different polyQ lengths. Furthermore, the validations performed in separate laboratories demonstrate the inter-laboratory performance and replicability of the assay. In this paper, we report the optimisation and adaptation of the previous assay procedure on the SMCxPRO[™] platform, along with two independent method validations according to international regulatory guidelines^{17,18}.

The key motivation for this study is the transition of the SMC CSF-based quantitation of mHTT levels from a research-grade assay to a clinical-grade assay. Specifically, previous methods have been fit to support exploratory measurements in the context of pre-clinical work and clinical trials (research grade) and have provided highly valuable scientific insight. The procedures and resulting data outlined here are fit for primary and secondary endpoint use in clinical trials (clinical grade) and may support the use of CSF mHTT data for registrational purposes during drug development^{17,18}.

Results

Assay optimisation. Capture antibody 2B7 was originally labelled with biotin and detection antibody MW1 was labelled with the fluorescent dye Alexa Fluor[®] 647 according to SMCxPRO[™] labelling kit manufacturer instructions, leading to high variability in assay performance. Optimisation of purification and labelling protocols together with thorough analytical characterisation of starting materials and end products improved performance and batch consistency of the assay reagents (Supplemental Table 1). Biotinylated antibody 2B7 showed purities >99% and a biotin incorporation rate of around 0.7, whereby the low value is favourable in preventing the formation of bead cross-links or aggregates. Labelling ratios ranging from 1:3.5 to 1:7.5 were tested for preparation of the Alexa Fluor[®] 647-labelled MW1 antibody. The highest labelling ratio delivered the highest incorporation rate and the highest fluorescence emission (Supplemental Table 1). A labelling ratio of 1:8 was used for production batches leading to an Alexa Fluor[®] 647 incorporation rate of 5.9 and a response/background signal ratio of >6 (Supplemental Table 1).

Suitability of the surrogate matrix was demonstrated during validation by parallelism experiments in patient samples (Table 4 and Supplemental Table 2). Due to difficulty in obtaining large amounts of human CSF from healthy donors, artificial CSF (aCSF) stabilised with 1% Tween20 and a protease inhibitor (cOmplete Protease Inhibitor[™] Cocktail) was used as a surrogate matrix for the preparation of calibration standard and quality control (QC) samples. Suitability of the surrogate matrix was investigated by comparing the performance of the calibration curve in pooled human CSF and the surrogate matrix (Supplemental Fig. 1). Using optimised capture and detection reagents together with the controlled assay matrix enhanced the assay signal-to-noise ratio at the lowest calibration (1.63 pg/mL HTT Q46) from 2–3 to 4–8 (assay pre-validation data, Supplemental Fig. 2). An improved signal-to-noise ratio was confirmed during assay validation (Supplemental Fig. 2).

After optimisation of pipetting and washing steps, an overall plate precision of the assay signals at a low QC (LQC) level of around 20% was reached and no systematic plate effects were observed. To mitigate low plate

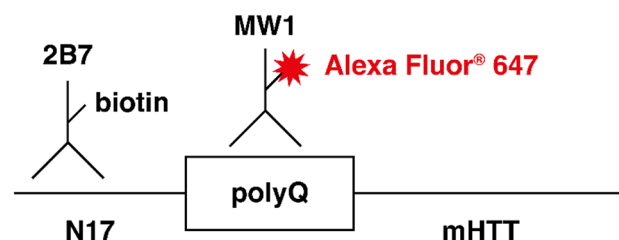


Figure 1. mHTT bead-based ligand binding assay: capture and detection of antibody-binding regions. mHTT, mutant huntingtin protein; polyQ, polyglutamine.

precision, all samples were analysed in triplicates to allow single-value outlier exclusion according to pre-defined criteria.

Assessment of assay specificity and the adequacy of the reference standard. *Measuring assay signals across a wide range of different recombinant HTT proteins.* To assess the impact of HTT size or polyQ-repeat length on the assay signal, the concentration–response curves of 21 different recombinant HTTs were measured (Fig. 2a). Assay signals across plates were normalised using repeated measures with HTT Q45, which had highly reproducible concentration-responses (Fig. 2b).

A graded increase in the assay response was observed with increasing polyQ repeats and flat responses among the proteins with low wild-type levels of polyQ repeats (cf. Q16 and Q23 proteins, Fig. 2a). As the number of

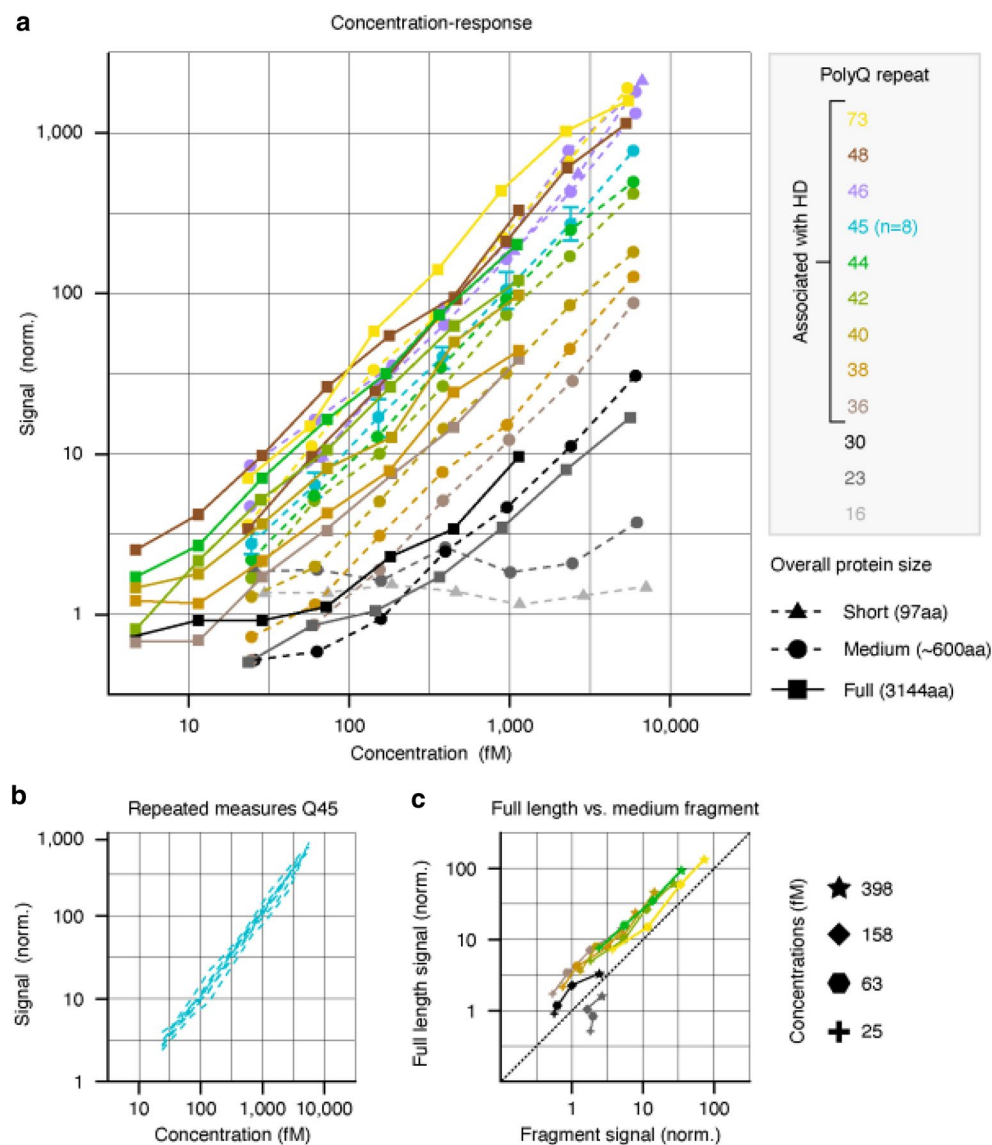


Figure 2. Concentration–response of different recombinant HTT. (a) Normalised signals across recombinant proteins varying in the number of polyQ repeats (12 repeats tested) and overall protein size (three sizes tested). The number of polyQ repeats is indicated by colour while overall protein size is indicated by marker symbols. The Q45 protein was measured across all eight plates and the highest concentration was used to derive a normalisation factor applied to all concentration-responses on the plate for cross-plate comparability. The error bars for the Q45 protein reflect the standard deviation across all eight runs. The Q46 medium size fragment was chosen as the reference protein during assay validation; both representative concentration-responses depicted in the figure were measured separately. (b) All individual concentration-responses of the Q45 protein measured across different plates. Note that normalisation induced the identical response at the highest concentration. (c) Normalised signals for fragment and full-length versions of proteins with the same polyQ repeat numbers and similar molarity. The same colour coding as for (a) was used to label the different polyQ repeat numbers. HD, Huntington’s disease; HTT, huntingtin protein; norm., normalisation; polyQ, polyglutamine.

polyQ repeats increased, an increase in assay response was observed, leading to robust concentration-responses that were particularly notable for recombinant proteins with polyQ-repeat lengths associated with HD (> 36). Among all proteins with ≥ 36 polyQ repeats, parallel, albeit shifted, concentration-responses were observed via log–log visualisation (Fig. 2a). Similar polyQ-repeat-dependent concentration-responses were observed for proteins of varying overall protein sizes (Supplemental Fig. 3). Wild-type polyQ-repeat lengths produced the lowest signal and flat responses; increasing polyQ repeats led to increased assay signal with overall parallel concentration-responses. Interestingly, we observed a partial overlap of assay signal for proteins with mid-40 polyQ-repeat lengths and the full-length Q73 protein, implying a potential saturation of signal towards high polyQ repeats. Such a potential signal saturation cannot be fully demonstrated with the current data and requires additional future investigation.

Next, a direct comparison of the assay signal was performed between medium-fragment and full-length proteins where the factors molarity and polyQ length could be kept identical (Fig. 2c). Note that the data for this direct comparison represent the subset of the data shown in Fig. 2a for which no other factors apart from protein size were different between the proteins. A three-way analysis of variance (ANOVA) of this subset of data confirmed that concentration, overall protein size and polyQ-repeat length all had significant main effects on the assay signal (all $P < 10^{-10}$, ANOVA), meaning each factor independently modulated the assay signal. Consistently higher signals were generated by the full-length proteins when compared with the medium-sized fragments ($P < 10^{-4}$, signed rank test of the signal difference between full-length proteins and medium-sized fragments across all polyQ-repeat lengths and concentrations). In addition, longer polyQ repeats led to consistent increases in the assay signal. Interestingly, the lowest polyQ-repeat length of 23 did not show a stronger signal for full-length protein when compared with fragment versions of the protein at the same molarity. However, a concentration–response (i.e. increasing signal with increasing concentration) was present for the full-length but not the fragment protein. These observations suggest that cross-reactivity between MW1 and wtHTT can occur at extremely high, likely non-physiological protein concentrations and is facilitated by the presence of full-length HTT.

Quantitative model comparison for the concentration–response curves. To quantitatively study whether concentration-responses for all recombinant proteins with ≥ 36 polyQ repeats were indeed parallel, two different linear mixed-effect models of the data were compared.

The model comparison aimed to assess whether a model that assumes the concentration-responses only differ by a constant offset between proteins (*Offset model*) performs equally well when compared with a model that assumes all proteins vary by a constant offset in addition to individual slopes for each protein (*Offset + Slope model*). Concentration-responses that were not parallel were only accommodated by the *Offset + Slope model*, while the *Offset model* implicitly assumed an identical slope for each protein, i.e. perfectly parallel concentration-responses. Importantly quantitatively parallel concentration-responses were to be indicated by comparative performances between the *Offset model* and the *Offset + Slope model*.

Data for the modelling were limited to proteins with longer polyQ repeats and to the signal within the linear working range of the assay, well within assay limits. Specifically, focus was placed on proteins with polyQ-repeat lengths ≥ 36 and concentrations > 26 femtomolar (fM) (i.e. above the lower limit of quantification [LLOQ]).

The *Offset model* (i.e. only vertical shifts in concentration–response curves with identical slopes) delivered essentially equivalent predictions ($r > 0.99$) compared with the *Offset + Slope model* (i.e. vertical shifts and individual slopes of the concentration–response for each protein) (Fig. 3a,b,c).

The nested nature of the models, i.e. whereby one model is identical to the other apart from one additional parameter, allowed for a formal nested model comparison. The nested model comparison revealed that the *Offset + Slope model* fitted the data significantly better compared with the *Offset model*, albeit with minor differences in actual model performance as assessed by the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) (log-likelihood ratio test, $P = 9 \times 10^{-4}$; *Offset model*: BIC = -193.8 , AIC = -205.7 ; *Offset + Slope model*: BIC = -197.8 , AIC = -215.7). In line with these results, we observed that the individual slopes of the *Offset + Slope model* only represented a minor (typically $< 5\%$) variation around a slope common to all proteins, consistent with the observation that the individual slopes had a minimal impact on the overall model's performance (Fig. 3c).

Another important question was whether the individual slopes of the *Offset + Slope model* that had a minor impact on overall model performance, but led to measurable improvements in the predictions, were completely random per protein or whether they systematically varied with protein properties such as the polyQ-repeat length. The individual slope estimates that were obtained in the *Offset + Slope model* correlated significantly and positively (Spearman rank correlation 0.73 , $P = 3.4 \times 10^{-5}$) with polyQ repeats. This observation suggested slightly steeper concentration-responses for proteins with higher polyQ-repeat numbers. Of note, the correlation between polyQ-repeat length and the individual slopes may potentially be driven by proteins with the lowest < 40 and highest > 70 repeats, while proteins within the mid-40 polyQ-repeat range varied more randomly (both positively and negatively) around the slope common to all proteins (e.g., see slopes for Q48 and Q42) (Fig. 4). Overall, the quantitative modelling revealed that the *Offset* and *Offset + Slope models* performed almost equivalently with small-but-measurable systematic individual response elements for different HTTs. These observations are in line with the idea that for practical purposes of assay conduct, the concentration–response curves across mutant versions of the recombinant HTTs are parallel.

Simulating longitudinal clinical visits to reveal the role of the reference standard. Given the above, the impact of specific recombinant proteins as a reference standard on actual assay results was investigated. Additionally, concentration estimates between patients with differing polyQ-repeat lengths were also compared. To address

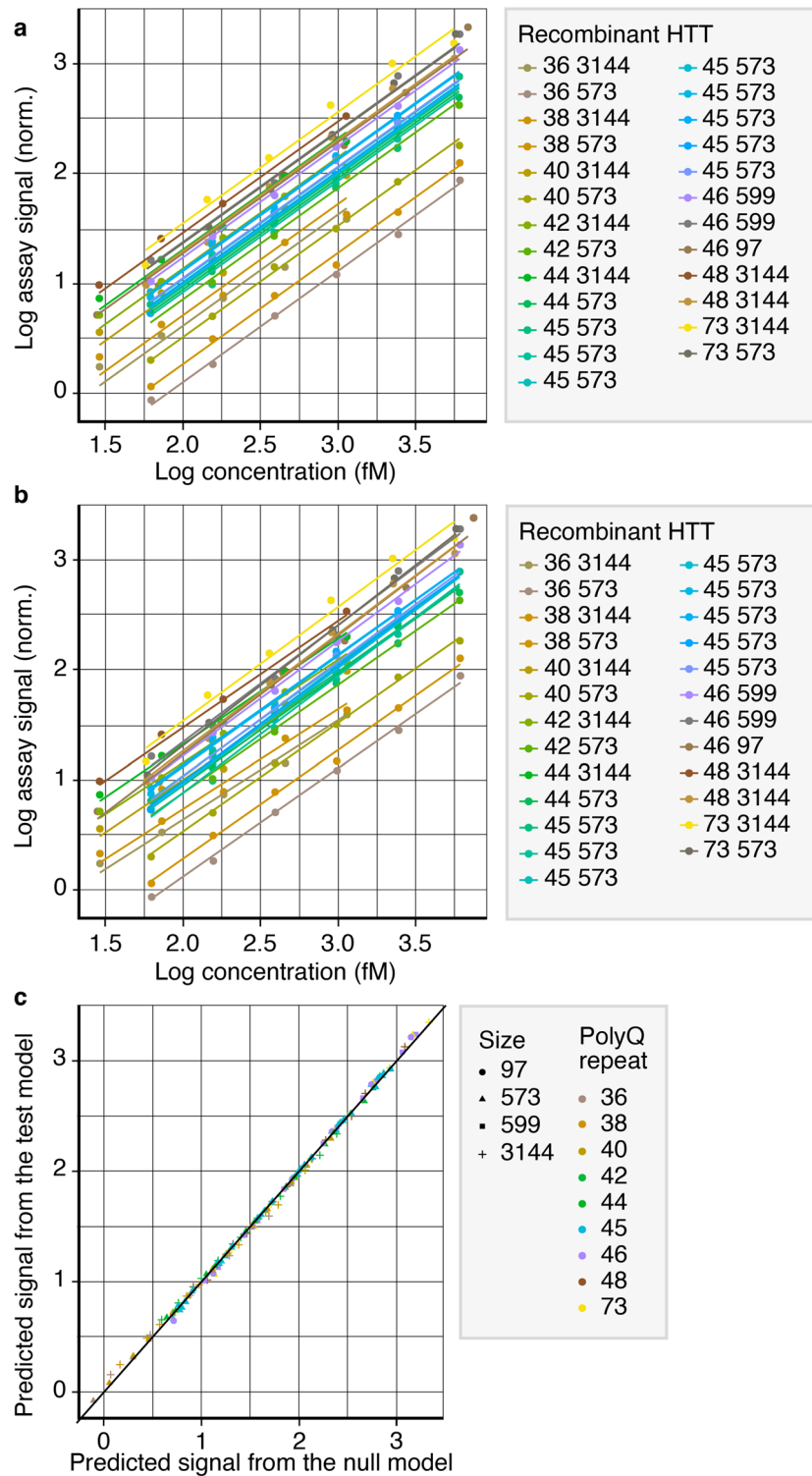


Figure 3. Model predictions of the *Offset* and *Offset+Slope* models across all concentration–response curves. (a) The *Offset* model was defined by random intercepts and fixed slopes. (b) The *Offset+Slope* model was defined by random intercepts and random slopes. In (a) and (b), the legend indicates each recombinant HTTs’ polyQ-repeat length, overall length and plate number, respectively. Legends in (a) and (b) indicate each recombinant HTTs’ polyQ-repeat length (first number) and protein size/number of amino acids (second number). (c) Comparison of predicted signals from the models in (a) and (b). Legend in (c) indicates the recombinant HTTs: icon shape denotes the protein size/number of amino acids; icon colour denotes the polyQ-repeat length. HTT, huntingtin protein; norm., normalisation; polyQ, polyglutamine.

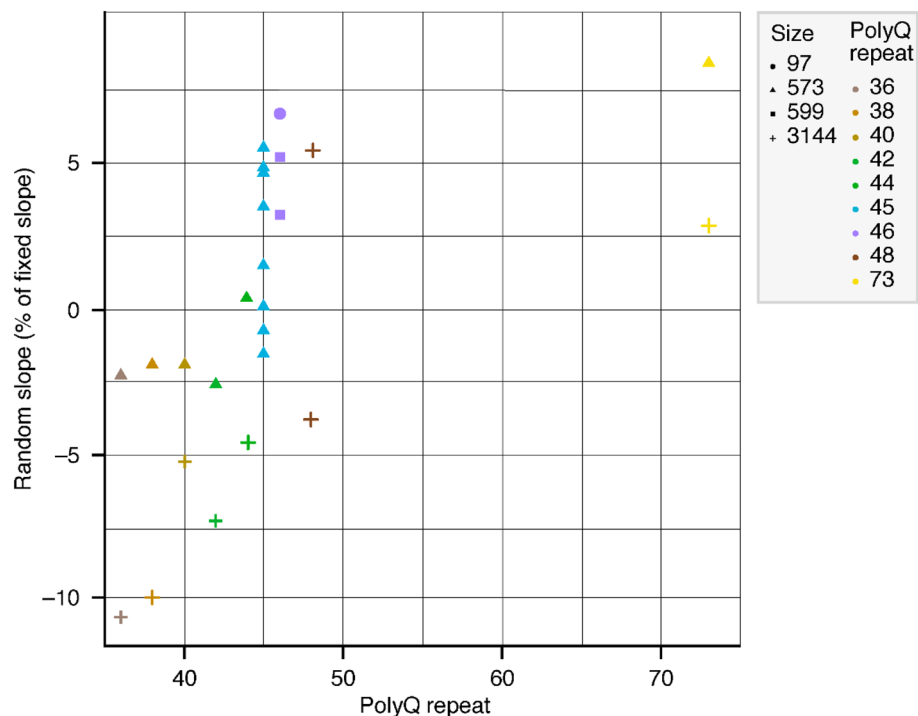


Figure 4. Relationship between polyQ-repeat length of recombinant HTT and random slopes. Legends indicate the recombinant HTTs: icon shape denotes the protein size/number of amino acids; icon colour denotes the polyQ-repeat length. HTT, huntingtin protein; polyQ, polyglutamine.

these questions and to assess the effect of the protein properties on the relative quantitation of CSF mHTT in a potential clinical trial setting, a simple and illustrative simulation was performed. Specifically, the *Offset model* and the *Offset + Slope model* were used to estimate concentrations for identical assay signals in a hypothetical longitudinal example.

A longitudinal mHTT-lowering signal, such as that potentially resulting from a drug-induced HTT-lowering approach, was simulated by assuming two assay signals at a putative baseline (Log signal = 1) and at a follow-up visit (Log signal = 0.7) (Figs. 2a and 5a–d).

The *Offset* and *Offset + Slope models* allowed for back calculating concentration estimates using the concentration–responses of each tested protein as a calibration curve. This allowed comparison of concentration estimates, both under the assumption that proteins only differ by a signal offset (i.e. response curves are fully parallel) and that proteins also differ by the individual slope as estimated in the *Offset + Slope model* above (i.e. response curves that systematically vary to a small degree with polyQ-repeat length).

In both the *Offset* and *Offset + Slope models*, the offsets between the concentration–response curves represented the largest source of variability in the resulting concentration estimates. Specifically, a reference protein with a higher overall concentration–response (vertical shift upward) would yield lower concentration estimates than a protein with a lower offset (vertical shift downward). The differences in concentration estimates were even on the order of several hundred fM for the identical assay signals. These observations emphasise the relative quantitative nature of the assay and the critical role of the reference standard in specifically influencing absolute concentration estimates (Fig. 5a,c). Given these results, historical data, generated with similar methods and reported as absolute concentrations, may need to be interpreted with caution, because the particular reference standard used in combination with the patient-specific polyQ-repeat region in the sample will render the absolute concentration estimates inaccurate.

Consistent with the above, investigations into longitudinal signals showed that the *Offset model* yielded widely different absolute change signals but identical percent change signals across the different recombinant HTTs (Fig. 5b). The identical percent change signal was a trivial consequence of the identical slope for each protein. This means that baseline and follow-up visit concentration estimates scaled identically for every protein, keeping the ratio between visits constant irrespective of the reference protein. In contrast, due to the variable slopes, the *Offset + Slope model* shows variations in both absolute and percent change across proteins (Fig. 5d). The percent change signal variations were typically $\pm 3\%$, which represented a minor fraction of the underlying simulated mHTT-lowering signal of approximately 50%.

Next, the simulation was expanded to allow for a wide range of longitudinal changes, including decreases from baseline as well as increases from baseline (Fig. 6). To this end, the simulated mHTT concentration change between baseline and follow-up visits was systematically varied in a putatively physiologically meaningful range (~60% decrease to 60% increase) between visits. The variability in the percentage change estimates between proteins was a function of the underlying mean percent change signal itself, whereby smaller percent changes

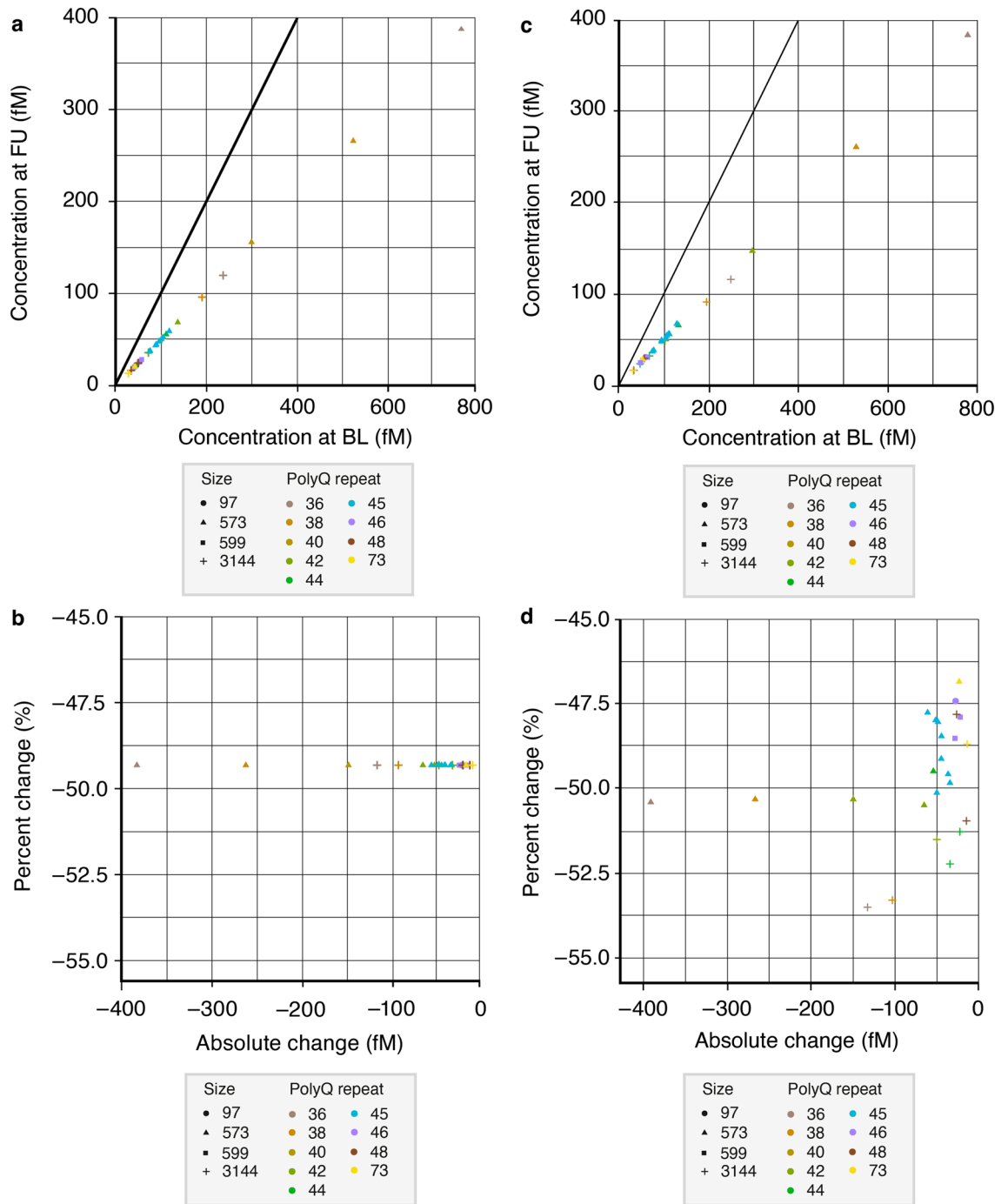


Figure 5. Simulation of longitudinal signals with an HTT-lowering effect. Two signal levels were used to simulate visits at baseline (Log signal = 1) and at follow-up (Log signal = 0.7), respectively. **(a)** The *Offset model* was defined by the random intercepts and fixed slopes that emerged from the *Offset model* fit to the concentration–response curve data. **(b)** Percent change plotted against absolute change in the *Offset model*. **(c)** The *Offset + Slope model* was defined by the random intercepts and random slopes that emerged from the *Offset + Slope model* fit to the concentration–response curve data. **(d)** Percent change plotted against absolute change in the *Offset + Slope model*. Legends outline the recombinant HTTs: icon shape denotes the protein size/number of amino acids; icon colour denotes the polyQ-repeat length. BL, baseline; FU, follow-up; HTT, huntingtin protein; polyQ, polyglutamine.

in mHTT were associated with less variability across reference standards. This means the greater the mean fold change signal (decrease or increase), the higher the resulting variability of the fold change estimates derived from the different recombinant HTTs under the *Offset + Slope model*.

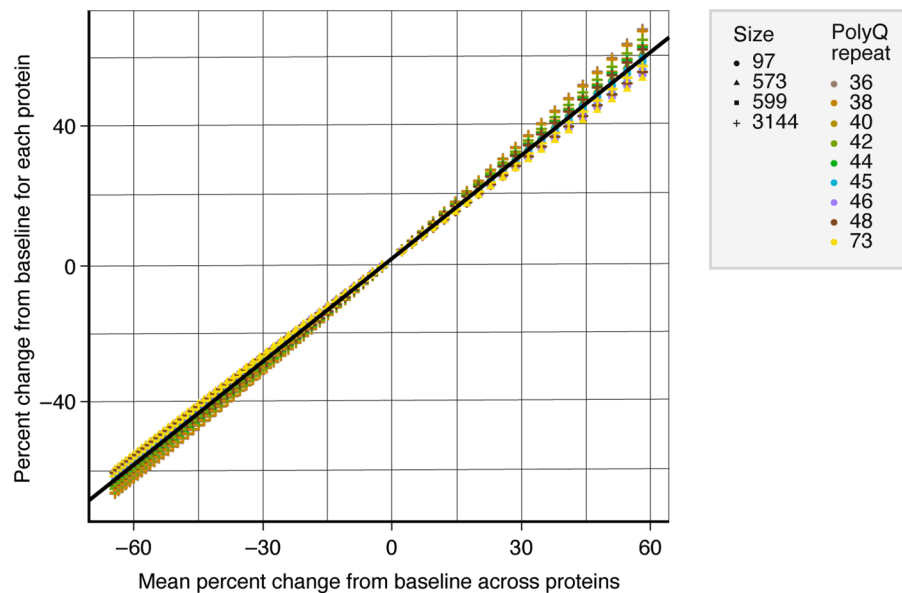


Figure 6. Variability of the estimated percent change signal in mHTT concentrations across recombinant proteins. Percent change from baseline for each recombinant HTT plotted against mean percent change from baseline across proteins. Legends indicate the recombinant HTTs: icon shape denotes the protein size (number of amino acids); icon colour denotes the polyQ-repeat length. HTT, huntingtin protein; mHTT, mutant HTT; polyQ, polyglutamine.

When the variability in the percentage change estimates (standard deviation across proteins) was normalised by the underlying signal change (mean across proteins) to derive the coefficient of variation, the fold change variability across recombinant HTTs constituted a minor fraction of the underlying change signal (in the order of 5%).

Overall, our observations suggest that for identical signals, there is major variability in the absolute concentration estimates and absolute change signals when different recombinant HTTs are used as a reference standard or when patients with differing polyQ-repeats are sampled in a trial. Importantly, this variability will render any determination of absolute concentrations inaccurate. Such inaccuracies may be present for historical data generated with similar methods and reported as absolute numbers. In contrast, relative change (percent change) signals are nearly identical across all reference proteins and a typical variation of 5% of the true change signal may be expected across the reference proteins tested. Such variability in relative change is likely negligible compared with physiological signals of interest and may not represent a confound for interpretation. Given these observations, any mid-40 polyQ-repeat HTT may be suitable as a reference standard for the assay and deliver equivalent relative quantitation of mHTT in CSF for clinical trials in adult-onset HD (polyQ-repeat range most consistently assessed here). The Q46 medium size fragment was used for assay validation, in keeping with previous research-grade versions of the assay. We expect the choice of a different reference standard with similar properties (e.g. a Q42 fragment or Q48 full length) to alter absolute concentration estimates, but to deliver highly consistent patterns of relative change.

Similarly, highly similar relative quantitation should emerge across patients with different polyQ-repeat lengths even if the particular reference standard used should not represent the polyQ length of that patient.

Different reference standards will therefore deliver comparable patterns of mHTT change across patients with a range of differing polyQ-repeat lengths. An important caveat associated with the different proteins is that for patients with very low numbers of polyQ repeats, it may empirically be more difficult to detect HTT in the CSF due to the lower overall assay signal delivered by low polyQ-repeat HTTs. In other words, assay sensitivity is likely to be a more pronounced problem for samples from patients within the reduced penetrance polyQ-repeat length range. Another more general limitation of the current data is with respect to juvenile HD, where the polyQ-repeat expansions can be larger than 100 repeats. Very high polyQ-repeats that are outside of the currently tested spectrum need further investigation to confirm whether parallel concentration-responses continue to be present.

Assay validation. *Performance of HTT Q46 calibrators during method validation.* Validation in two independent laboratories confirmed the assay had high sensitivity. The calibration range was 1.63 pg/mL (LLOQ—Roche) and 0.655 pg/mL (anchor point—ICON [validated LLOQ: 1.64 pg/mL]), to 400 pg/mL (upper limit of quantification [ULOQ]) HTT Q46 in surrogate matrix, determined via parallelism data. HTT Q46 calibrators prepared with reference standard spiked in surrogate matrix performed well during method validations (Table 1). Accuracy of all individual calibration samples were within the acceptance criteria of 70–130% accuracy. The calibrators enabled full recovery of frozen QC samples.

Intra- and inter-assay accuracy and precision in spiked surrogate matrix. Overall, inter- and intra-assay accuracy and precision of the reference standard (HTT Q46) in surrogate matrix (aCSF) matched the predefined acceptance criteria across both laboratories (Table 2). In both validations, the determined mean concentration at each level (including the LLOQ and ULOQ) was within 70–130% accuracy; precision of the mean concentration determined at each level was $\leq 30\%$ coefficient of variation (CV).

Inter-assay precision in the CSF of patients with HD. Precision of CSF samples from patients with HD were reliable across multiple independent runs within each laboratory as well as across the independent laboratories (Table 3). In both validations, the precision of the mean concentration determined for each patient sample met the acceptance criteria of $\leq 30\%$ CV precision.

Parallelism. To confirm comparable behaviour of different HTTs in patient CSF as the real matrix of interest, the assay signal of real patient CSF with largely differing polyQ-repeat lengths were measured (tested polyQ-repeat lengths: 41, 42, 44, 48, 50, 51) in serial dilution with surrogate matrix using the Q46 reference standard to back-calculate concentration estimates. Dilutional parallelism was demonstrated in all samples, including patients with low polyQ-repeat lengths (polyQ repeats 41, 42 (Supplemental Table 2); polyQ repeats 44–51 (Table 4)), further indicating a comparable assay signal behaviour across the spectrum of HD-related HTTs. The precision of the mean concentration across all dilutions within the dynamic assay range was $\leq 16.7\%$ across both validations, fulfilling the acceptance criteria of $\leq 30\%$ CV (Table 4). The parallelism data (Table 4) were used to determine the parallelism LLOQ (LLOQ_p). Comparable values of 1.57 pg/mL (Roche) and 1.69 pg/mL (ICON) were obtained.

Microplate homogeneity. Appropriate microplate homogeneity was demonstrated at the LLOQ level since the accuracies for all mimicry LLOQ samples were between 72.5 and 122.1% (Supplemental Table 3).

Nominal HTT Q46 concentrations in surrogate matrix (pg/mL)														
	Roche	ICON	Roche	ICON	Roche	ICON	Roche	ICON	Roche	ICON	Roche	ICON	Roche	ICON
	1.63	1.64	4.08	4.10	10.2	10.2	25.6	25.6	64.0	64.0	160	160	400	400
Mean	1.62	1.67	4.16	4.07	10.3	10.3	24.8	25.2	65.5	67.1	161	156	402	409
Precision (%CV)	1.5	6.5	4.3	4.7	3.0	4.3	4.6	4.5	4.8	6.9	6.0	6.9	4.4	6.8
Relative error /bias (%)	-0.6	1.7	2.0	-0.7	1.0	1.0	-3.1	-1.4	2.3	4.8	0.6	-2.6	0.5	2.3
n	18	24	18	25	18	25	18	25	18	25	18	25	18	25

Table 1. Precision and accuracy of calibration standards. Back-calculated mHTT concentrations (pg/mL) for calibration standards in surrogate matrix. CV coefficient of variation, HTT huntingtin protein, mHTT mutant HTT.

	LLOQ (pg/mL)		LQC (pg/mL)		MQC (pg/mL)		HQC (pg/mL)		ULOQ (pg/mL)	
	Roche	ICON	Roche	ICON	Roche	ICON	Roche	ICON	Roche	ICON
	1.63	1.64	4.50	4.50	40.0	40.0	300	300	400	400
Intra-assay										
Intra-assay accuracy (%)	84.7	100.3–119.7	71.8	100.2–115.5	79.0	97.4–120.2	73.3	80.5–115.8	79.3	80.8–120.0
Intra-assay precision (%CV)	14.1	14.1	12.1	5.6	6.2	4.0	4.1	6.2	6.7	12.9
Total error* (%)	29.4	NA	40.3	NA	27.2	NA	30.8	NA	27.4	NA
Inter-assay										
Inter-assay accuracy (%)	93.3	109.5	77.6	106.2	84.8	106.9	84.0	99.3	83.8	93.3
Inter-assay precision (%CV)	11.5	0.00	9.5	4.7	5.4	7.5	9.4	11.1	7.8	13.3
Total error* (%)	18.2	NA	31.9	NA	20.6	NA	25.4	NA	24.0	NA
Intra-assay and Inter-assay										
Overall bias (%)	NA	9.5	NA	6.2	NA	6.9	NA	-0.7	NA	-6.7
Total precision (%CV)	NA	14.1	NA	7.3	NA	8.5	NA	12.7	NA	18.5
Total error** (%)	NA	23.6	NA	13.5	NA	15.4	NA	13.4	NA	25.2

Table 2. Intra- and inter-assay accuracy and precision in spiked surrogate matrix. CV coefficient of variation, HQC high quality control, LLOQ lower limit of quantification, LQC lower quality control, MQC mid quality control, NA not applicable, ULOQ upper limit of quantification. *Sum of absolute value of bias and precision. **Sum of absolute value of overall bias and total precision.

Run ID	Found mHTT concentration (pg/mL) in patient CSF									
	Sample 1		Sample 2		Sample 3		Sample 4		Sample 5	
	Roche	ICON	Roche	ICON	Roche	ICON	Roche	ICON	Roche	ICON
Run 1	1.98	3.36	2.60	3.16	3.92	4.26	3.42	5.73	6.83	7.22
Run 2	2.05	2.55	2.55	3.03	3.75	3.83	3.25	3.90	6.28	7.21
Run 3	2.18	3.02	2.82	3.20	4.21	4.18	4.20	4.19	6.40	6.22
n	3	3	3	3	3	3	3	3	3	3
Mean conc. (pg/mL)	2.07	2.98	2.66	3.13	3.96	4.09	3.62	4.61	6.50	6.88
Precision (%CV)	4.9	13.7	5.4	2.9	5.9	5.5	14.0	21.4	4.4	8.3

Table 3. Inter-assay precision in the CSF of patients with HD. *Conc.* Concentration, *CSF* cerebrospinal fluid, *CV* coefficient of variation, *HD* Huntington's disease, *mHTT* mutant huntingtin protein.

Interferences of wtHTT, drug and blood. No relevant interferences of wtHTT (Supplemental Table 4) or drug (Supplemental Table 5) were shown in the assay. Interference of blood differed between the two independent validations. One laboratory observed interference at 1.0% whole blood in samples analysed and spiked at 1.64 pg/mL. In contrast, data from the other independent laboratory fulfilled the acceptance criteria for the absence of interference of blood (Supplemental Table 6).

Prozone effect. No high-concentration hook effect was observed up to the highest-tested concentration of 12.5 µg/mL of tominersen. The 1/100,000 dilution generated a result within the working range with an accuracy of 91.2%.

Stability of reference standard in surrogate matrix. Bench-top stability of the reference standard in surrogate matrix was demonstrated for at least 4 h (Supplemental Table 7). Stability of the reference standard in surrogate matrix was demonstrated for up to 12 months of storage at –60 °C to –85 °C (Supplemental Table 8) and for up to 85 days at –70 °C at ICON (Supplemental Table 9).

Incurred sample stability (ISS). The assessment of analyte stability in CSF was performed on patient samples. Multiple overlapping time periods were used to cover a study sample storage time of >3 years (Supplemental Fig. 4 and Supplemental Table 10). Time periods 1 and 2 were successfully validated, which led to a maximum demonstrated stability for study samples stored at –70 °C of 952 days (2.6 years).

Discussion

With the ongoing development of HTT-lowering therapies for HD, changes in the levels of CSF mHTT may be a critical biomarker that may capture a biological signal with direct causal relevance in the trajectory of HD pathology. To support regulatory decision-making processes in drug development, it is important to ensure biomarker assays are both robust and reliable. Furthermore, these assays should comply with international regulatory guidelines while maintaining transferability and generating replicable data.

The current study performed validations in two independent laboratories, aimed at generating a bead-based sandwich ligand binding assay that fulfils regulatory requirements as depicted in regulatory guidelines and is fit for primary and secondary endpoint use in clinical trials¹⁹. Additional analyses were also performed in this study to further characterise the assay methodology. Translation of the mHTT ligand binding assay from a research-grade environment to regulated validation in clinical-grade laboratories enables the assay to serve as a valuable resource that will facilitate the clinical development of HTT-lowering therapies.

Comparison of the assay signal across a wide range of recombinant HTTs showed a steep increase in assay signal, with the transition of HTT from non-disease causing (<36 polyQ repeats) to full penetrance (≥40 polyQ repeats). All tested recombinant mHTTs delivered robust concentration-responses that were highly parallel, supported by quantitative linear mixed-effect modelling data on multiple recombinant proteins with polyQ-repeat lengths ≥36. This was irrespective of polyQ length, overall protein size or expression system (Table 5). These observations suggest that the assay reported here is broadly relevant across the continuum of adult-onset HD, with comparable signal properties across a wide range of polyQ repeats.

Recent work in HD has shown the relevance of somatic expansion of the CAG tract in *HTT*²⁰. It is important to note that the somatic instability of *HTT* may likely lead to heterogeneous mixtures of polyQ-repeats being present in the CSF of patients. Also, given the dependency of the assay signal of repeat length, progressive expansion may lead to increasing longitudinal signals within long timescales and may confound estimates of changes in mHTT concentration. The data presented here suggest that the signal behaviour of polyQ repeats up to Q73 may be comparable. However, somatic expansion may lead to repeat lengths >100 for cells in some tissues. Very long repeat lengths currently remain a gap in our assessment. Information on the behaviour of polyQ repeats >100 would also be highly valuable in the context of juvenile HD.

Overall, several processes such as active and passive release of mHTT²¹, somatic expansion and pharmacological intervention²² may all impact the readout of mHTT in CSF, complicating the interpretation of mHTT

		Roche			ICON			
Sample	Dilution factor	Back-calculated conc. (pg/mL)	Undiluted conc. (pg/mL)	Recovery (%)	Back-calculated conc. (pg/mL)	Undiluted conc. (pg/mL)	Bias (%)	Cross-lab difference (%)
1	1	4.26	4.68	100	5.50	5.50	0	-16.1
	2	2.40	5.29	113	2.63	5.26	-4.3	0.6
	3	1.57*	5.19	110.9	1.64	4.91	-10.7	5.5
	4	1.56	6.87	NA	1.69*	6.78	23.2	
	6	0.601	3.96	84.6	NV	NA	NA	
	8	0.403	3.55	75.9	0.167	1.33	-75.7	
Mean (pg/mL)			5.05			5.61		10.5
CV (%)			6.5			14.5		
n (results within curve)			3			4		
2	1	5.07	5.57	100	6.96	6.96	0	-22.2
	2	2.53	5.57	100	3.45*	6.89	-1.0	-21.2
	3	1.79	5.90	105.9	1.60	4.79	-31.1	
	4	1.11	4.86	87.3	1.72	6.89	-0.9	
	6	0.762	5.03	90.3	1.02	6.11	-12.1	
	8	0.521*	4.58	82.2	0.625	5.00	-28.1	
Mean (pg/mL)			5.68			6.38		11.6
CV (%)			3.4			16.6		
n (results within curve)			3			4		
3	1	5.20	5.72	100	4.86	4.86	0	16.3
	2	2.45	5.40	94.4	1.97	3.95	-18.9	31.0
	3	1.55*, a	5.12	89.5	1.53	4.59	-5.7	10.9
	4	0.856	3.77	65.9	1.10	4.39	-9.7	
	6	0.749	4.94	86.4	0.646*	3.88	-20.3	
	8	0.454	3.99	69.8	0.803	6.43	32.1	
Mean (pg/mL)			5.41			4.47		-19.0
CV (%)			5.5			10.5		
n (results within curve)			3			3		
4	1	5.73	6.30	100	3.50	3.50	0	57.1
	2	2.83	6.23	98.9	1.60	3.20	-8.3	64.3
	3	2.20	7.27	115.4	1.01	3.04	-13.1	
	4	1.84	8.11	128.7	0.800*	3.20	-8.5	
	6	0.905	5.97	94.8	1.168	7.01	100.5	
	8	0.633*	5.57	88.4	0.410	3.28	-6.2	
Mean (pg/mL)			6.98			3.35		-70.3
CV (%)			12.8			6.1		
n (results within curve)			4			2		
5	1	8.58	9.44	100	NV	NA	NA	N/A
	2	4.21	9.26	98.1	2.69	5.38	0 [‡]	53.0
	3	2.57	8.48	89.8	1.77	5.32	-1.2 [‡]	45.8
	4	2.14	9.42	99.8	1.47	5.89	9.4 [‡]	46.1
	6	1.20	7.91	83.8	0.699*	4.20	-22.1 [‡]	
	8	1.05*	9.26	98.1	NV	NA	NA	
Mean (pg/mL)			9.15			5.53		-49.3
CV (%)			5.0			5.6		
n (results within curve)			4			3		
6	1	5.95	6.54	100	4.25	4.25	0	42.4
	2	2.73	6.01	91.9	1.91	3.81	-10.2	44.8
	3	NV	NA	NA	1.46*	4.39	3.3	N/A
	4	1.41 ^a	6.20	94.8	1.39	5.57	31.3	
	6	0.915	6.04	92.4	0.681	4.08	-3.8	
	8	0.574*	5.05	77.2	0.408	3.26	-23.1	
Mean (pg/mL)			6.25			4.50		-32.6
CV (%)			4.3			16.7		
n (results within curve)			3			4		

Table 4. Parallelism data for mHTT in the CSF of patients with HD. Serial dilution of samples from patients with mid–long polyQ-repeat lengths in surrogate matrix. *Lowest concentration for each sample providing a parallel response (i.e. recovery within 70–130%). #Because a triplicate CV > 20.0% was observed in the undiluted sample, the first twofold dilution result was used as a reference to calculate the bias. #Value just below the LLOQ_p of 1.57 pg/mL included in statistics. **Italic bold:** result outside 70–130% recovery, or absolute bias > 30%. Conc., concentration; CSF, cerebrospinal fluid; CV, coefficient of variation; HD, Huntington's disease; LLOQ_p, parallelism lower limit of quantification; mHTT, mutant huntingtin protein; NA, not applicable; NV, no value, duplicate well CV > 20%; polyQ, polyglutamine. Cross-lab difference (%) = (Result ICON—Result Roche) / (Result ICON + Result Roche)/2 * 100.

Fragment	Source	Expression system	Protein purification tag	Total protein size (amino acids)	Theoretical MW (kDa)	Purity by SDS-PAGE	Stock concentration (mg/mL)
HTT-Q16, 1–97	AMRI	BL21(DE3)	N-term MBP, C-term 6His	494	54.3	99.6%	1.00
HTT-Q23, 1–573	AMRI	SF9	N-terminal FLAG	579	63.0	98.0%	1.00
HTT-Q23, 1–3144	AMRI	HEK293	C-terminal FLAG	3162	350.0	96.0%	1.03
Q30-HTT (1–573)	AMRI	SF9	N-terminal FLAG	590	64.3	> 96.0%	1.00
Q30-HTT (1–3144)	AMRI	HEK293	C-terminal FLAG	3161	350.0	> 99.0%	1.06
Q36-HTT (1–573)	AMRI	SF9	N-terminal FLAG	596	65.1	> 98.0%	1.04
Q36-HTT (1–3144)	AMRI	HEK293	C-terminal FLAG	3167	351.0	> 98.0%	1.06
Q38-HTT (1–573)	AMRI	SF9	N-terminal FLAG	598	65.4	> 98.0%	0.95
Q38-HTT (1–3144)	AMRI	HEK293	C-terminal FLAG	3169	351.0	> 96.0%	1.07
Q40-HTT (1–573)	AMRI	SF9	N-terminal FLAG	600	65.6	> 97.0%	1.00
Q40-HTT (1–3144)	AMRI	HEK293	C-terminal FLAG	3171	351.0	> 97.0%	1.06
Q42-HTT (1–573)	AMRI	SF9	N-terminal FLAG	602	65.9	> 98.0%	1.01
Q42-HTT (1–3144)	AMRI	HEK293	C-terminal FLAG	3173	351.0	> 98.0%	1.04
Q44-HTT (1–573)	AMRI	SF9	N-terminal FLAG	604	66.1	> 98.0%	1.11
Q44-HTT (1–3144)	AMRI	HEK293	C-terminal FLAG	3175	352.0	> 99.0%	1.06
HTT-Q45 1–573	AMRI	SF9	N-terminal FLAG	601	65.8	99.0%	1.00
HTT-Q46, 599aa	Evotec	BL21(DE3)	N-terminal GPLGS	599	65.4	90.0%	0.25
HTT-Q46, 1–97	AMRI	BL21(DE3)	N-terminal MBP,				
C-terminal 6xHis	524	58.2	95.9%	0.76			
HTT-Q48, 1–3144	AMRI	HEK293	C-terminal FLAG	3187	353.0	95.2%	0.97
HTT-Q73, 1–573	AMRI	SF9	N-terminal FLAG	629	69.4	99.0%	1.00
HTT-Q73, 1–3144	AMRI	HEK293	C-terminal FLAG	3212	356.0	98.4%	1.11

Table 5. Overview of all HTT fragments. *His* histidine, *HTT* huntingtin protein, *MBP* maltose-binding protein, *MW* molecular weight, *Q* glutamine, *SDS-PAGE* sodium dodecyl-sulphate polyacrylamide gel electrophoresis.

levels as biomarker. It is an important task for the field to disentangle the individual mechanistic contributions of these processes on CSF mHTT readouts to improve our understanding of the clinical meaningfulness of changes.

The pre-set criteria for the validation parameters were met, fulfilling the requirements for precision and accuracy in spiked surrogate matrix; precision in CSF from individuals with HD; parallelism; specificity; prozone effect and microtiter plate homogeneity. Parallelism data also demonstrated the absence of a matrix effect, which refers to any impact of assay components, aside from the analyte, on the analytical properties of the assay²³. As parallelism experiments generate curves that represent the binding affinity of the analyte as well as interference from the matrix, the presence or absence of a matrix effect can be inferred from the success or failure of a parallelism experiment²⁴.

These validation findings support the reliability of this ultra-sensitive bioanalytical method for quantifying mHTT in human CSF and show that it can be replicated and transferred. Notably, given that this is now a state-of-the-art clinical-grade assay, previous findings on the levels of CSF mHTT in individuals with HD that were generated via the research-grade version of the assay may differ from future data.

An important limitation of the assay is its relative quantitative nature, requiring the choice of a particular reference standard against which heterogeneous patient samples are being compared. As a result, absolute concentrations that are estimated with this method may exhibit inter-patient variability of a technical rather than biological nature. In future experiments it will be important to disentangle technical and biological variance of the assay signal in greater detail cross-sectionally, to evaluate the assay's full scope as a biomarker tool. Despite the limitations of requiring a particular reference standard, absolute concentrations from this assay may potentially

carry valuable information about an individual's disease burden and trajectory when modelled appropriately in large data sets.

A further limitation is that sampling large amounts of CSF from healthy volunteers presents practical and ethical difficulties. As such, aCSF was used as a surrogate for human CSF during the preparation of calibrator and QC samples. The controlled environment afforded by the surrogate matrix is an important element for reliable assay performance, as justified by parallelism data.

Finally, the origin of blood interference in the bead-based assay has been attributed to clumping of magnetic particles as well as detection of mHTT present in the blood¹². The degree of interference in the assay may be influenced by the level of cell lysis achieved in the whole blood used as well as the amount of mHTT in the blood from donors. Validation experiments using blood from study participants in our assay are pending.

This bead-based sandwich ligand binding assay developed for the quantification of changes in mHTT levels in human CSF has been successfully characterised and independently validated in two laboratories. Our findings show that this assay may be a reliable tool for generating biomarker data in registrational clinical trials for HD, with relevance across the adult-onset HD continuum. Collaboration within the HD community will enable further refinement and application of this assay, supporting the development of HTT-lowering therapies for HD.

Methods

Materials. HTT fragments of different polyQ-repeat length and protein fragment size were purchased from AMRI and Evotec. Evotec HTT Q46, 599aa, was selected as the reference standard, and the surrogate matrix was purchased from Bio-Techne. Tween20 and protease inhibitor (cOmplete Protease Inhibitor™ Cocktail) were purchased from Sigma-Aldrich and Roche, respectively. The capture antibody was purchased from Evotec, and the detection antibody was expressed at Roche. SMC™ Capture Antibody Labeling kit (Merck Millipore) was used. Alexa Fluor® 647 carboxylic acid succinimidyl ester was purchased from Thermo Fisher Scientific. The centrifugal plate washer Blue® Washer (BlueCatBio, Germany) was used for plate wash steps. SMCxPRO™ specific buffers and glass-bottom 384-well plates were purchased from Merck Millipore. Ninety-six well polypropylene V-bottom plates were purchased from Brooks Life Sciences.

All human CSF samples were derived from individuals with early-manifest HD, obtained from the open-label extension (OLE) of the Phase I/IIa study of tominersen (NCT03342053).

Assay setup. All results were generated by Good Clinical Practice-trained personnel in a regulated bio-analytical environment with Good Laboratory Practice-certified laboratories. A bead-based sequential ligand binding assay with SMC detection was used on the SMCxPRO™ (Merck) platform.

The ultra-sensitive assay employs the antibody pair 2B7/MW1 for capture and detection (Fig. 1) and aCSF as a surrogate matrix. Capture antibody 2B7 binds to the N17 region of HTT (i.e. binds to both mHTT and wtHTT) and conjugates to streptavidin-coated magnetic particles via biotin coupling. Detection antibody MW1 is specific to the polyQ stretch present in mHTT and was labelled with Alexa Fluor® 647.

A 599 amino acid-long recombinant HTT fragment containing a Q46 amino acid-long polyQ chain was used as the reference standard (HTT Q46, molecular weight 65,390 g/mol). Calibration standard and QC samples were prepared in surrogate matrix containing 1% Tween20 and a protease inhibitor cocktail (surrogate matrix) for an assay range of 1.63–400 pg/mL. The minimum required dilution (MRD) of the assay was set to 2.

Optimisation of assay reagent preparation. Labelling of capture and detection antibodies were prepared at Roche Diagnostics (Penzberg, Germany), and labelled antibodies were analytically characterised at Roche Pharma Research (Penzberg, Germany). Antibody 2B7 was biotinylated using reagents and instructions from the SMC™ capture reagent labelling kit. The conjugate was purified via Superdex® 200 size exclusion chromatography. Coupling of biotinylated 2B7 to magnetic beads was performed according to kit instructions. MW1 was coupled with Alexa Fluor® 647-NHS ester with challenge ratios ranging from 1:3.5 to 1:8. Purification was performed via cut-off filtration (40 K MWCO Zeba™ Spin columns). Spectrophotometry was used to determine the antibody concentration and Alexa Fluor® 647 labelling rate. The purity was determined by size exclusion-high-performance liquid chromatography, and the fluorescence emission level was determined by fluorescence spectrophotometry. Labelled antibodies and coupled beads were stored at 2–8 °C.

Assay protocol. All calibration standards, QC samples and unknown samples were measured in triplicate wells. QC samples were prepared by spiking the reference standard in surrogate matrix, followed by shock freezing on dry ice and storage at or below –65 °C. Calibration samples were prepared on the day of the assay by spiking the reference standard in surrogate matrix. All buffers and reagents were equilibrated to room temperature before use and all assay steps were performed at room temperature.

A blocking buffer (50 µL/well) was dispensed into a 96-well V-bottom polypropylene plate, followed by either 150 µL/well calibration standards or a 15 µL/well sample dilution buffer containing 10% Tween20 and 1 × protease inhibitor, and a QC or study sample (135 µL/well). Coupled beads (100 µL/well) diluted 1:500 in an Assay Discovery Buffer were added to the plate using a 12-channel manual pipette. The plate was sealed and incubated for 1.5 h while shaken at 400 rpm. After incubation, the plate was placed for 2 min on the magnet of the Blue® Washer before centrifugation at 800 rpm. A sterile filtered detection antibody (20 µL/well) diluted 1:1000 in assay buffer was immediately added to the plate and incubated for 1 h in a shaker at 700 rpm. After incubation, the plate was placed for 2 min on the magnet of the Blue® Washer before performing four wash cycles at 800 rpm with a 200 µL System Buffer added at each step. The plate was incubated with the last wash buffer for 2 min in a shaker at 700 rpm and the solution was transferred to a second microplate using a 12-channel manual pipette. The plate was incubated for 2 min on the magnet of the Blue® Washer and the plate was centrifuged at 800 rpm

to remove the buffer. An elution buffer (12 $\mu\text{L}/\text{well}$) (Buffer B) was added to each well and the plate was placed in a shaker at 700 rpm for 6 min (performed using the Hamilton MicroLab Starline at ICON). A neutralisation buffer (10 $\mu\text{L}/\text{well}$) (Buffer D) was added to the glass-bottom 384-well reading plate using a manual 12-channel pipette (Roche) and the Hamilton MicroLab Starline (ICON). After placing the 96-well assay plate on a magnet for 2 min minimum, the supernatant (10 $\mu\text{L}/\text{well}$) was transferred from the 96-well assay plate to the 384-well reading plate using a manual 12-channel pipette (Roche) and the Hamilton MicroLab Starline (ICON). The 384-well reading plate was sealed with adhesive aluminium foil, placed in a shaker for 2 min at 700 rpm and spun for 5 min at 500 g. The plate was then placed on the bench for 30 min before readout on the SMCxPRO™ platform.

Data analysis was performed with Watson LIMS software (Roche) and Softmax Pro GxP 6.4 (ICON). For each triplicate well, the mean signal, standard deviation and the precision (%CV) were calculated. The calibration standards were fitted with a 4-parameter logistic with a weighting factor of $1/\text{mean signal}^2$. Concentrations of mHTT in samples were back-calculated using the fitted function and a minimum dilution factor of 1.1. Results of study samples showing signals below the LLOQ were reported as BLQ, provided they were measured at MRD.

Assay validation. Validation parameters and acceptance criteria were adapted to the context of use and to the assay performance observed during pre-validation experiments^{17,18}. All samples were analysed in triplicates and the mean assay signal reported if the triplicate precision was $\leq 20\%$. Due to the difficulty in obtaining sufficient human CSF from healthy donors, assay validations in both independent laboratories were performed using the surrogate matrix.

A calibration standard curve was developed, consisting of seven non-zero calibration standards covering the dynamic assay range (1.63/1.64–400 pg/mL). The LLOQ and ULOQ were defined as the lowest and highest calibration standard concentrations within the dynamic range, respectively. Acceptance criteria required a minimum of six non-zero calibrator levels to have an accuracy of 70–130%.

Inter-assay accuracy and precision in spiked surrogate matrix were assessed using a calibration standard curve and three sets of QC samples at the following concentrations: LQC (4.50 pg/mL), mid QC (40.0 pg/mL), high QC (HQC, 300 pg/mL) plus LLOQ/ULOQ samples. Intra-assay accuracy and precision were assessed using a calibration standard curve and three (ICON) or four (Roche) sets of QC samples at the five concentrations mentioned above for the inter-assay assessment. Measurements for inter- and intra-assay accuracy and precision were recorded in six independently prepared runs. Acceptance criteria for inter- and intra-assay accuracy and precision in spiked surrogate matrix required the determined mean concentration at each level including LLOQ and ULOQ to be within 70–130% accuracy; precision of the mean concentration determined at each level needed to be $\leq 30\%$ CV from the LLOQ to the ULOQ; and the total error needed to be $\leq 40\%$.

Inter-assay precision was also assessed for CSF samples from patients with HD, where five patient samples measured in three independently prepared runs were performed on three different days. Acceptance criteria required the precision of the mean concentration determined at each level to be $\leq 30\%$ CV from the LLOQ to the ULOQ.

Parallelism was assessed by the analysis of six samples from patients with HD. Study samples were serially diluted with surrogate matrix down to the LLOQ and below. Recoveries were calculated based on concentration of the sample diluted at the MRD. Acceptance criteria for parallelism experiments required precision of the mean concentration across all dilutions within the dynamic assay range to be $\leq 30\%$ CV for at least five out of the six tested samples. LLOQ_p was determined via parallelism data using the common concentration method²⁵ on data from six individuals with HD. Parallelism data were also used to validate the MRD.

Microplate homogeneity was assessed by adding a full set of calibration standards and QC samples, prepared in surrogate matrix, to an analytical run. A single volume of an LLOQ sample in surrogate matrix sufficient to fill all remaining free positions for validation samples of an analytical run (excluding calibration standards and analytical run acceptance QC samples) was prepared and added. The LLOQ sample was quantified using triplicate mean evaluation of study samples. Acceptance criteria for microplate homogeneity required $\geq 80\%$ of the LLOQ samples to show accuracies within 70% and 130%.

Potential interference of wtHTT was assessed using a full-length HTT containing a Q23 polyQ chain (HTT-Q23 1-3144 aa) spiked at 0, 20.0, and 200.0 pg/mL (corresponding to 0, 47.5 and 475.0 fM) in blank surrogate matrix and in surrogate matrix spiked at LLOQ and HQC levels. Interference of the drug on the assay was assessed at 0.0, 0.1, 1.0, 100.0 $\mu\text{g}/\text{mL}$ of tominersen in blank surrogate matrix and in surrogate matrix spiked at LLOQ and HQC levels. Interference of whole blood in the assay was assessed using increasing amounts of fully haemolysed whole blood from healthy volunteers (0, 0.1%, and 1.0% v/v) spiked in blank surrogate matrix and in surrogate matrix spiked at LLOQ and HQC levels. Acceptance criteria for the absence of interference at a given concentration of interfering compound required $\geq 66.7\%$ of the blank matrix aliquots (without reference standard) to have mean assay signals below LLOQ, and $\geq 66.7\%$ of the spiked matrix samples to show accuracies within 70% and 130%.

The prozone effect (high-dose hook effect) describes a phenomenon observed in sandwich immunoassays in which the assay signal becomes saturated and falls in the presence of very high analyte concentrations²⁶. The prozone effect was assessed by spiking surrogate matrix with the highest attainable reference standard concentration above the ULOQ (12.5 $\mu\text{g}/\text{mL}$). The final amount of surrogate matrix was $\geq 95\%$. The spiked sample was serially diluted in surrogate matrix to bring at least one concentration within the assay working range (1 analysis in triplicates per dilution factor), i.e. sample analysed undiluted, and diluted 1 in 10, 1 in 100, 1 in 1000, 1 in 10,000 and 1 in 100,000.

Stability of the reference standard in surrogate matrix was assessed at LQC and HQC levels. Samples were analysed in triplicate per concentration at the following conditions/time points: after 1 freeze/thaw cycle at $-60\text{ }^\circ\text{C}$ to $-85\text{ }^\circ\text{C}$; after 2 and 4 h at room temperature; after storage at $-60\text{ }^\circ\text{C}$ to $-85\text{ }^\circ\text{C}$ for approximately 1,

3, 6, and 12 months. Acceptance criteria required the accuracy of the mean concentration at each QC sample level to be within 70% and 130%; the precision of the mean concentration determined at each QC sample level needed to be $\leq 30\%$ CV; and a maximum of one result per set was allowed to be rejected for analytical reasons.

For the assessment of ISS, study samples stored at $-70\text{ }^{\circ}\text{C}$ and analysed within bioanalytical studies BN40423 or BN40955 at time point x (days from sample collection to first analysis) were reanalysed for ISS at time point y (days from sample collection to ISS analysis). Three different, overlapping x – y time periods were investigated using five samples for each x – y time period. The initial sample result at time point x was then compared with the ISS result at time point y (Supplemental Fig. 4). Acceptance criteria were set based on recommendations for ISS: bias reanalysis result compared with initial result $\leq 30.0\%$ for 2/3 of the samples (i.e., 10 out of 15 samples), with at least 3/5 samples at each time period.

Based on available assay development data, stability of the capture and detection antibodies mAb < mHTT > M-2B7-IgG-Bi lot BR02 and mAb < mHTT > M-MW1-IgG-Alexa647 lot BR08 was initially stated for 3 months of storage at 2 – $8\text{ }^{\circ}\text{C}$. The functional test consisting of the assessment of calibration curve performance was repeated after 3 months of storage at 2 – $8\text{ }^{\circ}\text{C}$. The stability of the capture and detection antibodies was considered acceptable if the signal-to-noise ratio at the LLOQ was ≥ 4 ; and the signal-to-noise ratio at the ULOQ was ≥ 1000 .

Characterising assay specificity for mHTT. All recombinant HTT concentration responses were measured in triplicate wells across a total of eight measurement plates. Triplicate precision values were typically below 25% (higher variability was observed for values near background signal). To compare responses across plates, the Q45 protein served as a reference signal that was measured on every plate. That means that the concentration–response of the Q45 protein is the only concentration–response that has been measured multiple times. Measuring the Q45 allowed the normalisation of all plate responses according to the signal at the highest concentration of the Q45 protein. Parallelism was tested on two pre-dose CSF samples from patients with polyQ-repeat lengths of 41 and 42 (clinical study ISIS 443139-CS2), in addition to parallelism experiments conducted during assay validation.

Relative concentration-responses were measured for recombinant HTT that varied in overall protein size, expression systems, vendors and polyQ-repeat numbers (Table 5). Assay signals were normalised using the HTT Q45 protein and protein concentrations expressed in fM.

To investigate suitability of the HTT Q46 reference standard for mHTTs with shorter polyQ-repeat lengths, parallelism was tested on two CSF samples prior to study drug injection, from patients with polyQ-repeat lengths of 41 and 42 (OLE of the Phase I/IIa study of tominersen). Study samples were serially diluted with surrogate matrix down to the LLOQ and below, with at least three different dilutions within the assay dynamic range and two below (e.g. MRD and additional 1:2, 1:3, 1:4, 1:6 and 1:8 dilutions). The recoveries were calculated based on concentration of the sample diluted at MRD.

Statistics. All modelling analyses for the comparison of concentration-responses curves across different recombinant HTT proteins were performed using RStudio v1.4.1717-3. Linear mixed effects models were fit using the lmer function from the lme4 package.

The *Offset model* with a random intercept for each unique recombinant protein was specified as:

$$S \sim C + (1|P)$$

where S is the measured assay signal, C is the nominal concentration and P is the unique recombinant proteins. This model was compared with the *Offset + Slope model* with random intercept and slope (nested model testing using ANOVA) which was specified as:

$$S \sim C + (1 + C|P)$$

Study approval. Samples were obtained from the OLE of the Phase I/IIa study of tominersen, which was approved by local ethics committees and conducted in accordance with the Declaration of Helsinki and the International Conference on Harmonization Guidelines for Good Clinical Practice.

Ethical approval and informed consent. The OLE study protocol was approved by the following ethics committees: NRES Committee London—West London and GTAC, London, UK; Ethik-Kommission der Medizinischen Fakultät der Universität Ulm, Germany; University of British Columbia Clinical Ethics Review Board, Canada. Informed consent was provided by all patients prior to participation in the OLE study.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 29 November 2022; Accepted: 30 March 2023

Published online: 01 April 2023

References

1. Roos, R. A. Huntington's disease: A clinical review. *Orphanet J. Rare Dis.* 5, 40. <https://doi.org/10.1186/1750-1172-5-40> (2010).

2. Ross, C. A. *et al.* Huntington disease: Natural history, biomarkers and prospects for therapeutics. *Nat. Rev. Neurol.* **10**, 204–216. <https://doi.org/10.1038/nrneuro.2014.24> (2014).
3. Tabrizi, S. J. *et al.* A biological classification of Huntington's disease: The Integrated Staging System. *Lancet Neurol.* **21**, 632–644. [https://doi.org/10.1016/S1474-4422\(22\)00120-X](https://doi.org/10.1016/S1474-4422(22)00120-X) (2022).
4. Tabrizi, S. J. *et al.* Potential disease-modifying therapies for Huntington's disease: Lessons learned and future opportunities. *Lancet Neurol.* **21**, 645–658. [https://doi.org/10.1016/s1474-4422\(22\)00121-1](https://doi.org/10.1016/s1474-4422(22)00121-1) (2022).
5. Aad, G. *et al.* Measurements of the nuclear modification factor for jets in Pb + Pb collisions at radical(s)NN] = 2.76 TeV with the ATLAS detector. *Phys. Rev. Lett.* **114**, 072302. <https://doi.org/10.1103/PhysRevLett.114.072302> (2015).
6. Potter, N. T., Spector, E. B. & Prior, T. W. Technical standards and guidelines for Huntington disease testing. *Genet. Med.* **6**, 61–65. <https://doi.org/10.1097/01.GIM.0000106165.74751.15> (2004).
7. The American College of Medical Genetics/American Society of Human Genetics Huntington Disease Genetic Testing Working Group. Laboratory guidelines for Huntington disease genetic testing. *Am. J. Hum. Genet.* **62**, 1243–1247 (1998).
8. Saudou, F. & Humbert, S. The biology of Huntingtin. *Neuron* **89**, 910–926. <https://doi.org/10.1016/j.neuron.2016.02.003> (2016).
9. Wild, E. J. *et al.* Quantification of mutant huntingtin protein in cerebrospinal fluid from Huntington's disease patients. *J. Clin. Invest.* **125**, 1979–1986. <https://doi.org/10.1172/JCI80743> (2015).
10. Byrne, L. M. *et al.* Evaluation of mutant huntingtin and neurofilament proteins as potential markers in Huntington's disease. *Sci. Transl. Med.* **10**, eaat7108 (2018).
11. Southwell, A. L. *et al.* Ultrasensitive measurement of huntingtin protein in cerebrospinal fluid demonstrates increase with Huntington disease stage and decrease following brain huntingtin suppression. *Sci. Rep.* **5**, 12166. <https://doi.org/10.1038/srep12166> (2015).
12. Fodale, V. *et al.* Validation of ultrasensitive mutant huntingtin detection in human cerebrospinal fluid by single molecule counting immunoassay. *J. Huntingt. Dis.* **6**, 349–361. <https://doi.org/10.3233/JHD-170269> (2017).
13. Tabrizi, S. J., Ghosh, R. & Leavitt, B. R. Huntingtin lowering strategies for disease modification in Huntington's disease. *Neuron* **101**, 801–819. <https://doi.org/10.1016/j.neuron.2019.01.039> (2019).
14. Ko, J., Ou, S. & Patterson, P. H. New anti-huntingtin monoclonal antibodies: Implications for huntingtin conformation and its binding proteins. *Brain Res. Bull.* **56**, 319–329. [https://doi.org/10.1016/s0361-9230\(01\)00599-8](https://doi.org/10.1016/s0361-9230(01)00599-8) (2001).
15. Aviolat, H. *et al.* Assessing average somatic CAG repeat instability at the protein level. *Sci. Rep.* **9**, 19152. <https://doi.org/10.1038/s41598-019-55202-x> (2019).
16. Li, P. *et al.* The structure of a polyQ-anti-polyQ complex reveals binding according to a linear lattice model. *Nat. Struct. Mol. Biol.* **14**, 381–387. <https://doi.org/10.1038/nsmb1234> (2007).
17. U.S. Dept. of Health and Human Services, F., CDER, CVM. *Guidance for Industry: Bioanalytical Method Validation.* (2018).
18. Piccoli, S. P., Sauer, J., Ackermann, B. *et al.* Points to consider document: Scientific and regulatory considerations for the analytical validation of assays used in the qualification of biomarkers in biological matrices. *Biomarker Assay Collaborative Evidentiary Considerations Writing Group, Critical Path Institute.* (2019).
19. World Health Organization. *UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases. Handbook: Good Laboratory Practice: Quality Practices for Regulated Non-clinical Research and Development.* (2001).
20. Kacher, R. *et al.* Propensity for somatic expansion increases over the course of life in Huntington disease. *Elife* <https://doi.org/10.7554/eLife.64674> (2021).
21. Caron, N. S. *et al.* Mutant Huntingtin is cleared from the brain via active mechanisms in Huntington disease. *J. Neurosci.* **41**, 780–796. <https://doi.org/10.1523/JNEUROSCI.1865-20.2020> (2021).
22. Tabrizi, S. J. *et al.* Targeting Huntingtin expression in patients with Huntington's disease. *N. Engl. J. Med.* **380**, 2307–2316. <https://doi.org/10.1056/NEJMoa1900907> (2019).
23. Guilbault, G. G. & Hjelm, M. International union of pure and applied chemistry, commission on analytical nomenclature and clinical chemistry division. Nomenclature for automated and mechanised analysis, (recommendations 1989). *Eur. J. Clin. Chem. Clin. Biochem.* **29**, 577–585. <https://doi.org/10.1351/pac198961091657> (1991).
24. Tu, J. & Bennett, P. Parallelism experiments to evaluate matrix effects, selectivity and sensitivity in ligand-binding assay method development: Pros and cons. *Bioanalysis* **9**, 1107–1122. <https://doi.org/10.4155/bio-2017-0084> (2017).
25. Stevenson, L. F. & Purushothama, S. Parallelism: Considerations for the development, validation and implementation of PK and biomarker ligand-binding assays. *Bioanalysis* **6**, 185–198. <https://doi.org/10.4155/bio.13.292> (2014).
26. Amarasiri Fernando, S. & Wilson, G. S. Studies of the 'hook' effect in the one-step sandwich immunoassay. *J. Immunol. Methods* **151**, 47–66. [https://doi.org/10.1016/0022-1759\(92\)90104-2](https://doi.org/10.1016/0022-1759(92)90104-2) (1992).
27. DeTora, L. M. *et al.* Good publication practice (GPP) guidelines for company-sponsored biomedical research: 2022 Update. *Ann. Intern. Med.* **175**, 1298–1304. <https://doi.org/10.7326/M22-1460> (2022).

Acknowledgements

The authors thank Shuang Song, of Chrysalis Medical Communications, UK, for providing editorial support for this manuscript, which was funded by F. Hoffmann-La Roche Ltd in accordance with Good Publication Practice (GPP 2022) guidelines²⁷. The authors would like to thank Bernhard Röttig and Tobias Oelschlägel for the optimisation and preparation of labelled reagents in this study; Daphne Adriaanse, Jaap Lubbers, Eline Sportel and Jessica de Ruiter for sample processing; and Jörg Hipp and Denis Engemann for their helpful comments and suggestions for statistical analyses. The authors would also like to thank Douglas Macdonald and the CHDI Foundation for the continued support and consultancy enabling the assay work presented here.

Author contributions

All authors contributed to the ongoing review and interpretation of the incoming data to inform the continued validation and finalisation of the assay. S.V., B.M., and D.J.H. designed the assay specificity experiments. S.V., M.M., E.S. and L.B. designed and refined the assay strategy and validation experiments. K.S. and A.B. conducted the assay development and validation experiments. S.V., B.M. and D.J.H. analysed the data and wrote the manuscript. D.J.H. performed statistical evaluation of the assay specificity data. N.G. supervised assay validation at ICON plc.

Funding

This study was funded by F. Hoffmann-La Roche Ltd. The OLE of the Phase I/IIa study of tominersen was initially sponsored by Ionis Pharmaceuticals, and transferred to F. Hoffmann-La Roche Ltd. SV, KS, BM, MM, AB, ES, LB and DJH are full-time employees of F. Hoffmann-La Roche Ltd. NG is a full-time employee of ICON Bioanalytical Laboratories.

Competing interests

SV, KS, BM, MM, AB, ES, LB and DJH are full-time employees of F. Hoffmann-La Roche Ltd. NG is a full-time employee of ICON Bioanalytical Laboratories.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-32630-4>.

Correspondence and requests for materials should be addressed to D.J.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023