

Communication

# The Association between CAG Repeat Length and Age of Onset of Juvenile-Onset Huntington's Disease

Jordan L. Schultz<sup>1,2,\*</sup>, Amelia D. Moser<sup>3</sup> and Peg C. Nopoulos<sup>1,2,4</sup> 

<sup>1</sup> Department of Psychiatry, Carver College of Medicine at the University of Iowa, Iowa City, IA 52242, USA; peggy-nopoulos@uiowa.edu

<sup>2</sup> Department of Neurology, Carver College of Medicine at the University of Iowa, Iowa City, IA 52242, USA

<sup>3</sup> Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO 80309, USA; amelia.moser@colorado.edu

<sup>4</sup> Department of Pediatrics, Carver College of Medicine at the University of Iowa, Iowa City, IA 52242, USA

\* Correspondence: jordan-schultz@uiowa.edu; Tel.: +1-319-384-9388

Received: 27 July 2020; Accepted: 18 August 2020; Published: 20 August 2020



**Abstract:** There is a known negative association between cytosine–adenine–guanine (CAG) repeat length and the age of motor onset (AMO) in adult-onset Huntington's Disease (AOHD). This relationship is less clear in patients with juvenile-onset Huntington's disease (JOHD), however, given the rarity of this patient population. The aim of this study was to investigate this relationship amongst a relatively large group of patients with JOHD using data from the Kids-JOHD study. Additionally, we analyzed data from the Enroll-HD platform and the Predict-HD study to compare the relationship between CAG repeat length and AMO amongst patients with AOHD to that amongst patients with JOHD using linear regression models. In line with previous reports, the variance in AMO that was predicted by CAG repeat length was 59% ( $p < 0.0001$ ) in the Predict-HD study and 57% from the Enroll-HD platform ( $p < 0.0001$ ). However, CAG repeat length predicted 84% of the variance in AMO amongst participants from the Kids-JOHD study ( $p < 0.0001$ ). These results indicate that there may be a stronger relationship between CAG repeat length and AMO in patients with JOHD as compared to patients with AOHD. These results provide additional information that may help to model disease progression of JOHD, which is beneficial for the planning and implementation of future clinical trials.

**Keywords:** CAG; juvenile-onset Huntington's disease; motor onset

## 1. Introduction

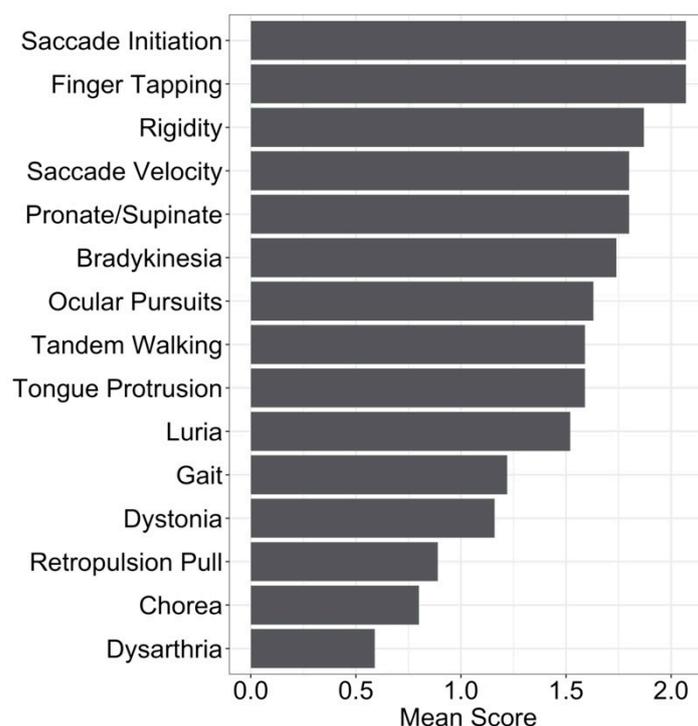
Huntington's disease (HD) is an autosomal dominant, neurodegenerative disorder that causes cognitive, behavioral, and motor symptoms [1]. An abnormal number of repeats ( $\geq 36$ ) of cytosine–adenine–guanine (CAG) within the huntingtin gene causes the mutation that leads to HD. The negative relationship between the number of CAG repeats that a person has and the age of motor onset (AMO) has been well-established for patients with adult-onset HD (AOHD) [2–6]. However, given the rarity of juvenile-onset HD (JOHD), the relationship between CAG repeat length and AMO is less clear. In general, it is believed that the correlation between CAG repeat length and AMO increases in higher CAG repeat lengths [7–9], but these reports are limited by small sample sizes. A previous review paper utilized data from 15 patients with JOHD who had a CAG repeat length of  $>60$  to demonstrate that the association between AMO and CAG seemed to be higher at CAG lengths of 60 to 80. However, at CAG repeat lengths above approximately 80, the association between AMO and CAG became much weaker [10]. In order to further investigate this phenomenon in a larger group, the same authors gathered patient data from seven separate case studies or case series of patients with JOHD and found 26 patients with a CAG repeat length of  $>80$ . In this cohort,

CAG repeat length accounted for only 26% of the variance in AMO [10]. A recent systematic review was performed and gathered information on more than 200 patients with reported JOHD [11]. This cohort of patients included a wide range of CAG repeats from 39 to 265. The authors reported a Pearson's correlation coefficient of  $-0.56$  between CAG repeat length and the age of onset of clinical symptoms. In line with other studies, the strength of correlation seemed to weaken at longer CAG repeats [11]. Both of these large reviews are limited, however, by their reliance on a collection of data, rather than primary collection from a single cohort. Thus, the aim of the present study was to investigate the relationship between CAG repeat length and AMO amongst patients with JOHD, with the goal of providing additional clarity to this current gap in knowledge. We also leveraged data from the Enroll-HD platform and the Predict-HD study to compare the results of patients with JOHD to those of patients with AOHD.

## 2. Materials and Methods

### 2.1. Description of Data

We analyzed data from a large sample of participants enrolled in the Kids-JOHD studies ( $N = 27$ ). These studies recruited participants from across the United States who were between the ages of 6 and 25 at the time of their first study visit and had a confirmed diagnosis of JOHD. To be enrolled, all participants were required to have already had molecular confirmation of HD and a clinical diagnosis of JOHD before traveling to the University of Iowa for assessment. To confirm that the participant was manifesting significant motor symptoms (meaning that they would meet criteria for a clinical diagnosis of JOHD), we utilized a combination of (1) detailed parental report of motor symptoms; (2) clinical evaluation by a pediatric neurologist at the University of Iowa; and (3) a Unified Huntington's Disease Ratings Scale (UHDRS) Total Motor Score (TMS) [12] of at least 12 to consider a participant to be motor-manifest. Each subject was required to have all three to be considered motor-manifest. Although a TMS of 12 may be considered low, patients with JOHD often present with far greater hypokinetic symptoms than do those with AOHD, and the UHDRS may not be sufficiently sensitive to these as it was designed for AOHD. These criteria excluded 10 individuals enrolled in the Kids-JOHD study and represented subjects who were tested locally mainly based on family history or behavioral issues, rather than significant motor symptoms. Of note, the TMS derived from the UHDRS takes into account various motor symptoms beyond just chorea. At their baseline visit, the participants with JOHD demonstrated the highest scores in saccade initiation, finger tapping, and rigidity (Figure 1). Additionally, the summative chorea score was one of the areas where patients with JOHD had the lowest score. This is of particular importance to consider when confirming the diagnosis of JOHD, as symptoms other than chorea can often be present. Participants were also excluded if they had a history of brain surgery or significant head trauma. All participants and their guardians (if under 18) signed informed consent prior to enrolling in these studies, which were approved by the University of Iowa Institutional Review Board (IRB).



**Figure 1.** Mean scores on the various domains of the Unified Huntington's Disease Rating Scale (UHDRS). Specifically, ocular pursuit score is the mean of horizontal and vertical pursuit scores; saccade initiation is the mean of the horizontal and vertical initiation scores; saccade velocity is the mean of the horizontal and vertical velocity scores; finger tapping is the mean of the right and left hand scores; pronate/supinate is the mean of the right and left hand scores; rigidity is the mean of the scores in the right and left arms; dystonia score is the mean scores from the trunk, right and left upper extremities, and right and left lower extremities; chorea score is the mean of maximal chorea scores from the face, buccal-oral-lingual scores, trunk, right and left upper extremities, and right and left lower extremities. All other scores include only one component and represent the mean of that domain.

We also analyzed data from the Predict-HD study [13] and from the Enroll-HD platform [14]. Specifically, we included participants from both studies who received a motor diagnosis of HD during the study after the age of 21 and who had a CAG repeat length of less than 60. This was done to ensure that we were investigating only those participants with adult-onset HD (AOHD). Participant reporting of historical motor onset timing is included in the Enroll-HD platform, but this may be subject to significant recall bias. Participants must have had a motor exam conducted with a diagnostic confidence level (DCL) of less than four to be included in the analyses from Enroll-HD and Predict-HD. The age at which these participants had their first report of a diagnostic confidence level of four on the UHDRS [12] was considered the age of motor onset (AMO). There were 242 participants who received a motor diagnosis during the Predict-HD study and 782 participants in the Enroll-HD database.

## 2.2. Statistical Analysis

The primary analysis investigated the relationship between CAG repeat length and AMO amongst the JOHD participants using simple regression models. We used natural cubic splines to transform the independent variable (CAG repeat length) to investigate nonlinear relationships for the primary analysis. RStudio was used for all analyses, with a  $p$ -value of  $<0.05$  considered statistically significant.

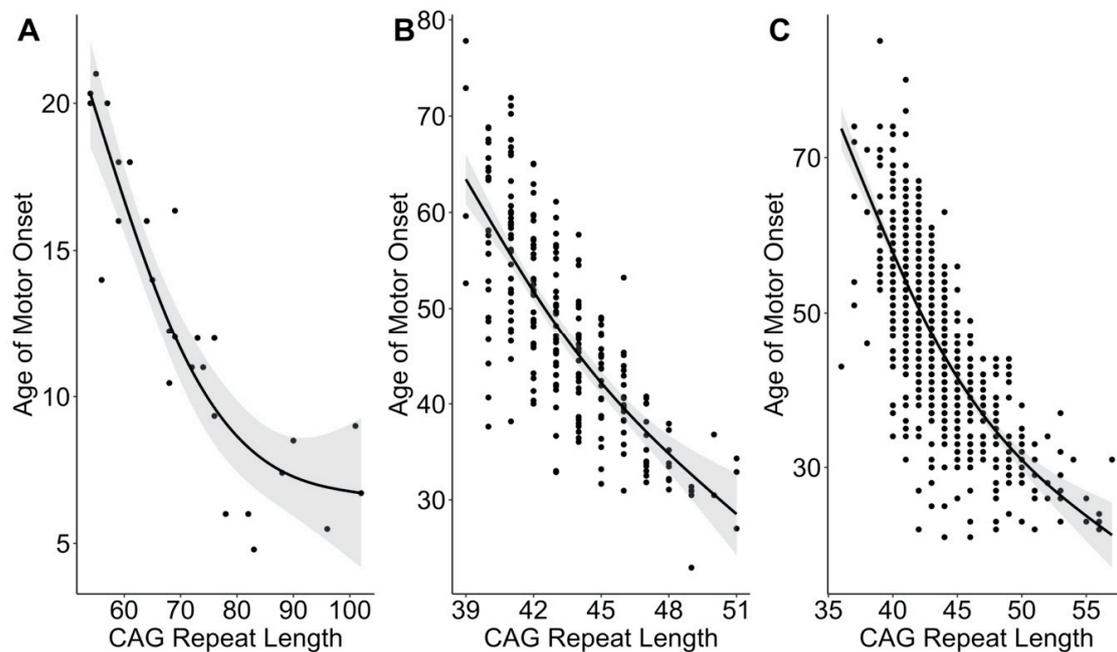
## 2.3. Institutional Ethical Approval

The University of Iowa institutional review board initially approved the study on 10/13/2011 (IRB # 201109879). For participants younger than 18 years, parents or guardians provided written

consent and children provided written assent. For participants who were 18 years or older, participants provided written consent.

### 3. Results

The participants from the Kids-JOHD study had a significant nonlinear relationship between AMO and CAG repeat length ( $R^2 = 0.84$ ,  $p = 2.63 \times 10^{-10}$ ) (Figure 2A). As has been reported previously, we also observed significant nonlinear relationships between AMO and CAG repeat length using data from participants with AOHD from the Predict-HD and Enroll-HD studies. From Predict-HD, the  $R^2$  value was 0.59 ( $p = 2.2 \times 10^{-16}$ ) (Figure 2B). The results from participants from Enroll-HD closely matched those observed in Predict-HD. Specifically, the  $R^2$  value was 0.57 ( $p = 2.2 \times 10^{-16}$ ) (Figure 2C).



**Figure 2.** Cytosine–adenine–guanine (CAG) repeat length significantly predicts the age of motor onset in (A) patients with juvenile-onset Huntington’s disease (JOHD) from the Kids-JOHD study, (B) patients with adult-onset Huntington’s Disease (AOHD) from the Predict-HD study, and (C) patients with AOHD from the Enroll-HD study. Black lines show the predicted regression lines, and the gray ribbons display the 95% confidence intervals.

### 4. Discussion

In the present analysis, we demonstrated that the CAG repeat length accounted for over 80% of the variance in AMO amongst patients with JOHD. This was substantially higher than in either of the AOHD analyses conducted, which showed CAG repeat length accounting for 59% of the variance (Predict-HD sample) and 57% of the variance (Enroll-HD). These results support previous reports demonstrating potentially increased predictive power of higher CAG repeat lengths [7–9]. Of note, previous reports have demonstrated that the predictive power of CAG repeat length on AMO seems to decrease at the highest CAG repeats of approximately 80 or above [10,11]. Our current cohort only included seven participants with a CAG repeat length above 80. Therefore, we did not have sufficient data to formally analyze whether or not the relationship between CAG and AMO weakens at higher CAG repeat lengths. However, informally, our results seem to confirm these previous reports. In Figure 2A, it seems as though there is a strong, linear relationship between CAG repeat length and AMO in those participants with a CAG repeat length of <80. There seems to be a bend in the regression line at approximately a CAG repeat length of 80, where the line begins to flatten out. This same shape was seen in previous reports with larger numbers of patients [10,11]. This may be due to a floor effect in the

ability of the CAG repeat length to predict AMO. Specifically, it may be possible that neurodegenerative changes occur over the course of 3–5 years. Therefore, the earliest possible AMO may be approximately five years old, regardless of CAG repeat length. This is only a hypothesis, though. Another possible explanation for the weakened relationship between CAG and AMO at CAG repeats above 80 may be related to the role that the huntingtin protein plays in neurodevelopment [15,16]. Neurodevelopmental changes have been reported to be more prominent at higher CAG repeats in patients with AOHD [15]. At higher CAG repeats, it is possible that the neurodevelopmental aberrations play a major role in the AMO in addition to neurodegeneration. This likely leads to significant difficulty in determining when their actual AMO is, likely resulting in a significant amount of heterogeneity in age of diagnosis.

One possible explanation for why higher CAG repeat lengths (>60) may explain more of the variance in AMO is that longer CAG repeat lengths may play a greater role in the development of pathologic changes that impact the onset of disease. Another consideration is that it is known that environmental exposures may modify disease onset in patients with AOHD [17–19]. Patients with JOHD may not have the same opportunity to be exposed to particular environmental factors. Therefore, their AMO is more closely linked to CAG repeat length alone and not the additional impact of environmental factors. Genetic modifiers of disease onset have also been identified in patients with AOHD [20,21]. These studies rely on large numbers of patients to identify genetic modifiers that may impact disease onset in AOHD. It cannot be ruled out that specific genetic modifiers exist that could impact the AMO in JOHD that have not been identified, given the rarity of this patient population.

One of the largest previous reports of the relationship between CAG repeat length and AMO amongst patients with JOHD utilized data from the Italian Huntington's Disease Databank, which retrospectively collected data from patients at two separate institutions and only includes 15 patients with a CAG repeat length of >60 [10,22]. This same review gathered data from seven separate case reports and case series and identified 26 patients with more than 80 CAG repeats [10]. Given the means by which these data were collected, the conclusions drawn in the resulting review may be seen as preliminary, as they are not the result of primary data collection. However, using the present large dataset of patients with JOHD, we are now able to confirm these previous findings showing that the strength of correlation between CAG and AMO is greater in JOHD, predicting about 84% of the variance in AMO [10].

There are important limitations to this study. First, despite being one of the largest studies of JOHD in the world, the number of patients is still relatively small. Second, we did not implement natural logarithmic transformations of our data, which has been recommended in previous studies investigating CAG and AMO [23]. We opted to not use natural logarithmic transformations of the data because doing so appeared to lead to disproportionate variance across the groups, which would lead to heteroscedasticity in the data. Lastly, as mentioned previously, the diagnosis of motor symptoms in children with the longest CAG repeats can be quite difficult and subject to bias and variability.

## 5. Conclusions

This study has provided further confirmation that CAG repeat length significantly predicts AMO in patients with JOHD. In fact, more than 80% of the variance of AMO was explained by CAG repeat length. This provides additional information that allows for more accurate modeling of JOHD, which is critical for future clinical trials.

**Author Contributions:** Conceptualization, J.L.S., A.D.M., and P.C.N.; methodology, J.L.S. and P.C.N.; software, J.L.S.; validation, A.D.M. and P.C.N.; formal analysis, J.L.S.; investigation, J.L.S. and A.D.M.; resources, P.C.N.; data curation, P.C.N.; writing—original draft preparation, J.L.S.; writing—review and editing, A.D.M. and P.C.N.; visualization, J.L.S.; supervision, P.C.N.; project administration, P.C.N.; funding acquisition, P.C.N. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by National Institute of Neurological Disorders and Stroke (NINDS; R01NS055903). The APC was funded by National Institute of Neurological Disorders and Stroke (NINDS; R01NS055903).

**Acknowledgments:** Enroll-HD is a longitudinal observational study for Huntington’s disease families intended to accelerate progress towards therapeutics; it is sponsored by CHDI Foundation, a nonprofit biomedical research organization exclusively dedicated to developing therapeutics for HD. Enroll-HD would not be possible without the vital contributions of the research participants and their families. We also thank the patients and families that participated in the Predict-HD study and the Kids-JOHD study. We also wish to extend our thanks to the CHDI Foundation for their dedication to advancing scientific discovery in the HD community through data sharing.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Macdonald, M.E.; Ambrose, C.M.; Duyao, M.P.; Myers, R.H.; Lin, C.; Srinidhi, L.; Barnes, G.; Taylor, S.A.; James, M.; Groot, N.; et al. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington’s disease chromosomes. *Cell* **1993**, *72*, 971–983. [[CrossRef](#)]
2. Langbehn, D.R.; Brinkman, R.; Falush, D.; Paulsen, J.S.; Hayden, M.R. A new model for prediction of the age of onset and penetrance for Huntington’s disease based on CAG length. *Clin. Genet.* **2004**, *65*, 267–277. [[CrossRef](#)]
3. Stine, O.; Pleasant, N.; Franz, M.L.; Abbott, M.H.; Folstein, S.E.; Ross, C.A. Correlation between the onset age of Huntington’s disease and length of the trinucleotide repeat in IT-15. *Hum. Mol. Genet.* **1993**, *2*, 1547–1549. [[CrossRef](#)]
4. Lucotte, G.; Turpin, J.; Riess, O.; Epplen, J.T.; Siedlacz, I.; Loirat, F.; Hazout, S. Confidence intervals for predicted age of onset, given the size of (CAG) n repeat, in Huntington’s disease. *Qual. Life Res.* **1995**, *95*, 231–232. [[CrossRef](#)]
5. Andrew, S.E.; Goldberg, Y.P.; Kremer, B.; Telenius, H.; Theilmann, J.; Adam, S.; Starr, E.; Squitieri, F.; Lin, B.; Kalchman, M.A.; et al. The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington’s disease. *Nat. Genet.* **1993**, *4*, 398–403. [[CrossRef](#)]
6. Squitieri, F.; Sabbadini, G.; Mandich, P.; Gellera, C.; Di Maria, E.; Bellone, E.; Castellotti, B.; Nargi, E.; De Grazia, U.; Frontali, M.; et al. Family and molecular data for a fine analysis of age at onset in Huntington disease. *Am. J. Med Genet.* **2000**, *95*, 366–373. [[CrossRef](#)]
7. Metzger, S.; Rong, J.; Nguyen, H.H.P.; Cape, A.; Tomiuk, J.; Soehn, A.S.; Freudenberg-Hua, Y.; Propping, P.; Freudenberg, J.; Tong, L.; et al. Huntingtin-associated protein-1 is a modifier of the age-at-onset of Huntington’s disease. *Hum. Mol. Genet.* **2008**, *17*, 1137–1146. [[CrossRef](#)]
8. Sun, Y.-M.; Zhang, Y.-B.; Wu, Z.-Y. Huntington’s Disease: Relationship Between Phenotype and Genotype. *Mol. Neurobiol.* **2016**, *54*, 342–348. [[CrossRef](#)]
9. Chao, T.-K.; Hu, J.; Pringsheim, T. Risk factors for the onset and progression of Huntington disease. *NeuroToxicology* **2017**, *61*, 79–99. [[CrossRef](#)]
10. Squitieri, F.; Frati, L.; Ciarmiello, A.; Lastoria, S.; Quarrell, O. Juvenile Huntington’s disease: Does a dosage-effect pathogenic mechanism differ from the classical adult disease? *Mech. Ageing Dev.* **2006**, *127*, 208–212. [[CrossRef](#)]
11. Cronin, T.; Rosser, A.; Massey, T.H. Clinical Presentation and Features of Juvenile-Onset Huntington’s Disease: A Systematic Review. *J. Huntington’s Dis.* **2019**, *8*, 171–179. [[CrossRef](#)] [[PubMed](#)]
12. Kiebertz, K.; Penney, J.B.; Corno, P.; Ranen, N.; Shoulson, I.; Feigin, A.; Abwender, D.; Greenarnyre, J.T.; Higgins, D.; Marshall, F.J.; et al. Unified Huntington’s disease rating scale: Reliability and consistency. *Mov. Disord.* **1996**, *11*, 136–142. [[CrossRef](#)]
13. Paulsen, J.S.; Long, J.D.; Rosser, A.E.; Harrington, D.L.; Erwin, C.J.; Williams, J.K.; Westervelt, H.J.; Johnson, H.J.; Aylward, E.H.; Zhang, Y.; et al. Prediction of manifest Huntington’s disease with clinical and imaging measures: A prospective observational study. *Lancet Neurol.* **2014**, *13*, 1193–1201. [[CrossRef](#)]
14. Landwehrmeyer, G.B.; Fitzer-Attas, C.J.; Giuliano, J.D.; Gonçalves, N.; Anderson, K.E.; Cardoso, F.; Ferreira, J.J.; Mestre, T.A.; Stout, J.C.; Sampaio, C. Data Analytics from Enroll-HD, a Global Clinical Research Platform for Huntington’s Disease. *Mov. Disord. Clin. Pr.* **2016**, *4*, 212–224. [[CrossRef](#)] [[PubMed](#)]
15. Van Der Plas, E.; Langbehn, D.R.; Conrad, A.L.; Kosciak, T.R.; Tereshchenko, A.; Epping, E.; Magnotta, V.; Nopoulos, P. Abnormal brain development in child and adolescent carriers of mutant huntingtin. *Neurol.* **2019**, *93*, 1021–1030. [[CrossRef](#)]

16. Saudou, F.; Humbert, S. The Biology of Huntingtin. *Neuron* **2016**, *89*, 910–926. [[CrossRef](#)]
17. Schultz, J.; Kamholz, J.A.; Moser, D.J.; Feely, S.M.; Paulsen, J.S.; Nopoulos, P. Substance abuse may hasten motor onset of Huntington disease. *Neurol.* **2017**, *88*, 909–915. [[CrossRef](#)]
18. Schultz, J.L.; Nopoulos, P.C.; Killoran, A.; Kamholz, J.A. Statin use and delayed onset of Huntington’s disease. *Mov. Disord.* **2018**, *34*, 281–285. [[CrossRef](#)]
19. Schultz, J.; Harshman, L.A.; Langbehn, D.R.; Nopoulos, P.C. Hypertension Is Associated with an Earlier Age of Onset of Huntington’s Disease. *Mov. Disord.* **2020**. [[CrossRef](#)]
20. Lee, J.-M.; Wheeler, V.C.; Chao, M.J.; Vonsattel, J.P.G.; Pinto, R.M.; Lucente, D.; Abu-Elneel, K.; Ramos, E.M.; Mysore, J.S.; Gillis, T.; et al. Identification of Genetic Factors that Modify Clinical Onset of Huntington’s Disease. *Cell* **2015**, *162*, 516–526. [[CrossRef](#)]
21. Lee, J.-M.; Correia, K.; Loupe, J.; Kim, K.-H.; Barker, D.; Hong, E.P.; Chao, M.J.; Long, J.D.; Lucente, D.; Vonsattel, J.P.G.; et al. CAG Repeat Not Polyglutamine Length Determines Timing of Huntington’s Disease Onset. *Cell* **2019**, *178*, 887–900.e4. [[CrossRef](#)] [[PubMed](#)]
22. Squitieri, F.; Berardelli, A.; Nargi, E.; Castellotti, B.; Mariotti, C.; Cannella, M.; Lavitrano, M.L.; De Grazia, U.; Gellera, C.; Ruggieri, S. Atypical movement disorders in the early stages of Huntington’s disease: Clinical and genetic analysis. *Clin. Genet.* **2000**, *58*, 50–56. [[CrossRef](#)] [[PubMed](#)]
23. Langbehn, D.R.; Hayden, M.R.; Paulsen, J.S.; The PREDICT-HD Investigators of the Huntington Study Group. CAG-repeat length and the age of onset in Huntington disease (HD): A review and validation study of statistical approaches. *Am. J. Med Genet. Part B Neuropsychiatr. Genet.* **2009**, *153B*, 397–408. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).