



Published in final edited form as:

Am J Med Genet A. 2022 April ; 188(4): 1056–1064. doi:10.1002/ajmg.a.62607.

Use of the Vineland-3, a measure of adaptive functioning, in CLN3

An N. Dang Do, MD PhD¹, Audrey E. Thurm, PhD², Cristan A. Farmer, PhD², Ariane G. Soldatos, MD³, Colby E. Chlebowski, PhD², Julie K. O'Reilly, PsyD², Forbes D. Porter, MD PhD¹

¹*Eunice Kennedy Shriver National Institute of Child Health and Human Development, NIH, Bethesda, Maryland, USA*

²*National Institute of Mental Health, NIH, Bethesda, Maryland, USA*

³*National Institute of Neurological Disorders and Stroke, NIH, Bethesda, Maryland, USA*

Abstract

Progressive vision loss and neurocognitive impairment are early and frequent presentations in CLN3 disease. This highlights neurodevelopmental functioning as critical to the disease, but limits the neuropsychological test repertoire. We evaluated the convergent validity of the Vineland Adaptive Behavior Scales as a potential outcome measure. In a prospective observational study of 22 individuals (Female:Male 11:11; 6-20 years-old) with a molecular diagnosis of CLN3, we used generalized linear models and Spearman correlations to quantify the relationship of the adaptive behavior composite (ABC) standard score with established outcomes of verbal IQ (VIQ) and disease severity (Unified Batten Disease Rating Scale, UBDRS) scores. We analyzed ABC changes in 1-year follow-up data in a subset of the same cohort (n=17). The ABC and VIQ, both standard scores, exhibited a strong positive correlation in cross-sectional data ($r=0.81$). ABC and UBDRS scores were strongly and positively correlated in cross-sectional data ($r_{\text{range}}=0.87-0.93$). Participants' ABC scores decreased slightly over the 1-year follow-up period (mean change, 95% CI: $-5.23, -2.16$). The convergent validity of the Vineland-3 for use in CLN3 is supported by its relationships with the established outcomes of VIQ and UBDRS. Future longitudinal research, including replication in other cohorts and evaluation of sensitivity to change, will be important to establish utility of the Vineland-3 for monitoring change in CLN3.

Correspondence: An Ngoc Dang Do, MD PhD, 10 Center Drive, MSC 1103. Bethesda, MD 20892-1103, Phone: 301.496.8849. an.dangdo@nih.gov.

Authors Contributions

ADD, AET, FDP conceptualized the study. ADD, AET, AGS, CEC, JKO performed the study and collected the data. ADD, AET, CAF analyzed and interpreted the data. ADD, CAF, AET drafted the manuscript. ADD, AET and FDP supervised and secured funding. All authors revised and approved the manuscript for submission.

Competing interest statement: An N. Dang Do, Audrey Thurm, Cristan Farmer, Ariane Soldatos, Colby Chlebowski, Julie O'Reilly, Forbes D. Porter declare that they have no conflict of interest in connection with this article.

Ethics approval and patient consent

The Institutional Review Board of the *Eunice Kennedy Shriver National Institute of Child Health and Human Development* approved the protocol (Investigations of Juvenile Neuronal Ceroid Lipofuscinosis; [NCT03307304](#)). All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000. Informed consent was obtained from all participants prior to enrolling in the study.

Keywords

adaptive behavior; Batten; natural history

1 INTRODUCTION

CLN3 disease (OMIM 204200) is an autosomal recessive, neurodegenerative, lysosomal disorder resulting from pathogenic variants in *CLN3* (Anderson, Goebel, & Simonati, 2013; Mole & Cotman, 2015). The disease is marked by childhood onset of vision loss and progression to blindness, among other neurologic impairments (Adams, Mink, & University of Rochester Batten Center Study, 2013; Preising, Abura, Jager, Wassill, & Lorenz, 2017). While treatment development is underway (Johnson et al., 2019; Specchio et al., 2020), a limited number of outcome measures, particularly focused on functioning over time (*Assessing Neurocognitive Outcomes in Inborn Errors of Metabolism*, <https://www.fda.gov/media/96785/download>), constrains research of the natural history, as well as efficacy evaluations for potential therapies. Cognitive performance is a common outcome (Adams et al., 2013; Augustine et al., 2019; Lamminranta et al., 2001), but at least a portion of the IQ tests used to measure cognition are invalidated by vision loss. Frequently used IQ tests have decreased sensitivity in the lower range (Farmer et al., 2020), limiting their use in the setting of progressive neurocognitive decline in CLN3. Adaptive behavior, the collection of skills required for age-appropriate functioning in one's environment, is correlated but not redundant with cognitive performance (Bolte & Poustka, 2002; Freeman, Ritvo, Yokota, Childs, & Pollard, 1988; Hayes & Farnill, 2003; Keith, Fehrmann, Harrison, & Pottebaum, 1987). It is an important and real-world measure of neuropsychological functioning (Adams et al., 2006).

The use of adaptive behavior is particularly applicable for early onset neurodegenerative conditions (*Assessing Neurocognitive Outcomes in Inborn Errors of Metabolism*, <https://www.fda.gov/media/96785/download>), due to both clinical relevance and the psychometric properties of the instruments. Adaptive behavior deficits are a core symptom of intellectual disability, and so measures are developed to be valid for most individuals, regardless of physical disability or cognitive status. Few neurocognitive measures have been specifically designed and validated for individuals with vision loss and intellectual disability, an issue that limits their use in CLN3 disease. CLN3-specific test batteries emphasizing verbal modalities have been implemented (Adams et al., 2013; Lamminranta et al., 2001), but the progression of behavioral, speech, and neurocognitive dysfunction eventually renders them inappropriate. Even using only the verbal standardized cognitive assessments, children's scores will likely continue to diverge negatively from those of their peers and cluster at the lower, far less sensitive and reliable, end of the scale. For current therapeutic approaches aimed at slowing, rather than reversing, neurodegeneration, outcome measures that enable meaningful longitudinal comparisons of non-standardized scores (e.g., person ability (Farmer et al., 2020)) would be more applicable for trials.

Here, we evaluate a broadly used measure of adaptive behavior, the Vineland Adaptive Behavior Scales, 3rd Edition, for use as a trial outcome measure in CLN3-related disorders.

We focused on the convergent validity of the scale, operationalized as its relationship to established outcomes of verbal IQ and disease state. Because CLN3 is a neurodegenerative disorder, we also looked to the relationship of Vineland scores with age for evidence of validity. For validity to be supported, we expected to see that 1) lower adaptive behavior scores would be related to lower verbal IQ and more severe Unified Batten Disease Rating Scale scores and 2) adaptive behavior scores would be worse for older than younger classically presenting participants. Given that CLN3 is a degenerative condition, adaptive functioning should decrease over time, though relevant data from other outcomes suggest that disease progression over childhood may take years (Kuper et al., 2019; Lamminranta et al., 2001; Masten et al., 2020). This gave rise to a final, more exploratory expectation, that 3) within classically presenting participants, worsening in adaptive functioning would be observed over a one-year follow-up period.

2 PARTICIPANTS and METHODS

2.1 Ethics Statement

The Institutional Review Board of NICHD approved the protocol [Investigations of Juvenile Neuronal Ceroid Lipofuscinosis (CLN3); [NCT033007304](#)]. We enrolled individuals of any age or disease severity, with two *trans* variants in *CLN3*, clinical symptoms consistent with CLN3 disease [Juvenile Neuronal Ceroid Lipofuscinosis (JNCL); Batten disease; OMIM 204200], and ability to travel to the NIH Clinical Center without increased health risk. *CLN3* variants were determined from reports from CLIA-certified laboratories. We obtained consent from parents/guardians or individuals 18 years of age, and assent from individuals 7 years as developmentally appropriate. Participants from this study have a consistent identifier (SP_._.) across publications to facilitate data comparisons.

2.2 Disease Severity Assessment

We evaluated disease severity as measured by the Unified Batten Disease Rating Scale (UBDRS, version 12/20/17) (Marshall et al., 2005). In the current analyses, we used the UBDRS subdomain assessments of Physical (a 27-item exam of speech, vision, and musculoskeletal functions), Capability given actual vision (a 5-item interview of daily living abilities), and Clinical Global Impression of severity (CGI; a 7-item provider's rating of overall disease state). Higher Physical and CGI scores indicate more severe, whereas higher Capability scores associate with less severe impairment. Out of the six UBDRS subdomains, these three subdomains assess the actual overall disease state. Calculations of the weighted UBDRS scores and the reason and comparisons for using the 7-item versus the single overall CGI score are as previously described (Dang Do et al., 2021).

2.3 Neuropsychological Assessment

The neuropsychological evaluation utilized a tiered approach for IQ testing, as in other studies of CNS dysfunction (Dickson et al., 2011), including in this population. Either the *Wechsler Intelligence Scale for Children-Fifth Edition* (WISC-V) (Wechsler, 2014) or the *Wechsler Adult Intelligence Scale, Fourth Edition* (WAIS-IV) (Wechsler, 2008) was administered, based on age of the participant. Similar to previous studies of CLN3 (Adams et al., 2013; Lamminranta et al., 2001), verbal IQ (VIQ) is reported since visual impairment

precluded administration of the nonverbal IQ in many participants. VIQ estimates could not be obtained for three participants, due to language barriers (SP7.2.1 and SP11.2.4) or neurologic state (SP17.2.1). VIQ is a norm-referenced score with a population mean of 100 and standard deviation of 15 (range 45 – 155 on the WISC-V and 50-150 on the WAIS-IV).

We report here on standard scores from the Communication, Daily Living, and Socialization domains and the Adaptive Behavior Composite (ABC) score of the Vineland Adaptive Behavior Scales, Third Edition (Vineland-3) (Sparrow, Cicchetti, & Saulnier, 2016) semi-structured caregiver interview version. The norm-referenced standard scores have a population mean of 100, standard deviation of 15, and range of 20 – 160. We do not report standard scores for the motor domain, since they are not available for children older than 9 years. The sub-domains are scored using scaled (v-scale; mean=15, SD=3) or person ability (growth scale value, GSV; unitless) scores. Standard errors of the mean (SEMs) differ by age and subscale, but approximate values can be used to help interpret changes relative to measurement error: ABC (SEM~3), domain-level standard scores (SEM~4), subdomain V-scale scores (SEM~1.5), and GSV (~3). Since GSV scores are available regardless of age, Motor subdomain scores are included in GSV analyses.

Of note, the Vineland-3 manual reports limited normative data on a sample of individuals with visual impairments; the ABC scores in this reference sample (M=86.8, SD=20.4) were lower and more variable than the average reference sample (M=100, SD=15). While the Wechsler tests excluded people with uncorrected visual impairments from standardization samples, VIQ scores from these tests have been used in studies of CLN3 (Adams et al., 2013; Lamminranta et al., 2001). Data on children with only visual impairments suggest that they tend to score lower than their peers on some VIQ domains (i.e., comprehension, similarity, and vocabulary), but may perform relatively better on working memory subtests (Morash & McKerracher, 2017). Importantly, the Vineland-3 has a lower minimum value than the Wechsler scales (20 versus 45-50), which must be considered in any comparison between the two assessments.

2.4 Data Analysis

Descriptive statistics for the phenotypic variables are presented for the total sample and by genotype. We assessed the relationships among adaptive behavior (Vineland-3), cognitive ability (VIQ), and disease severity (UBDRS) separately for baseline and follow-up. Although correlation is the standard statistic associated with convergent validity, we initially planned a generalized linear model, given the presence of four sibling dyads (SP5.2.1 and 5.2.2; SP10.2.1 and 10.2.5; SP12.2.1 and 12.2.2; SP16.2.1 and 16.2.2) in the data and the possibility of non-linearity in the relationships. However, through exploratory data analysis we found that the random effect of family was zero in all models, so we ultimately used the DescTools package (Signorell & al., 2020) to calculate Spearman correlations and their 95% CI for all associations.

We also examined the reliability of the Vineland-3 ABC with VIQ, both of which are measured on the same scale, by calculating the concordance correlation coefficient. Values above 0.5 reflect at least some agreement, but to justify interchangeability, the CCC should

approach 1.0. The cccrm package (Carrasco, 2015) was used to estimate the CCC and the 95% CI.

We explored the cross-sectional effect of age using a generalized linear model with age as the lone fixed effect predicting Vineland score, and a random subject-level intercept to account for repeated measurement. Because the goal of this specific analysis was to produce evidence consistent with neurodegeneration, we excluded the two participants with vision-only presentation.

We also looked for evidence of neurodegeneration in the subset of participants with follow-up data (excluding those with vision-only presentation). We tested the hypothesis that change from baseline to follow-up (approximately 1 year) in Vineland scores would differ from zero, regardless of chronological age. This was accomplished using a generalized model with fixed effects of measurement occasion (baseline or follow-up) and exact time-to-follow-up, as well as a random subject-level intercept.

Through standard exploratory data analysis, we evaluated the tenability of the assumptions of linearity, homoscedasticity, and normality of the model residuals. Following the current guidelines of the American Statistical Association (Wasserstein, Schirm, & Lazar, 2019), we do not categorize p-values as “statistically significant” and instead focus on the magnitude and variability of the effects via parameter estimates and 95% CI. We used R version 4.0.0(R-Core-Team, 2020) for data analysis and visualization. The code for all analyses is available upon request.

3 RESULTS

From October 2017 – April 2019 we enrolled 22 participants with CLN3-related disorders from 18 families (i.e., four sibling pairs) (Supplemental Table). The cohort includes participants from 6 – 20 years old (mean = 11 years 7 months, SD = 4 years 4 months), of equal sex distribution (Female:Male 11:11). Participants reported white non-Hispanic (n=14), white Hispanic/Latino (n=3), multiple (n=3), and Asian non-Hispanic (n=2) race/ethnicity. Ten participants were homozygous and five were compound heterozygous for the common 1-kb deletion (c.461-280_677+382del) variant. The genotype and presentation of sibling pair SP5.2.1/5.2.2 were at the time consistent with a vision-only CLN3 phenotype.

Based on parental reports, all participants had visual impairments at baseline visit (mean age at symptom onset 6 years 5 months, SD = 1 years 11 months). Estimated visual acuity showed 20/22 qualifying for legal blindness (i.e., < 20/200) status. Other symptoms included caregiver reported changes in neurocognitive ability (n=17, onset mean=6 years 7 months, SD=4 years 4 months), behavioral or psychiatric symptoms (n=21, onset mean=7 years 5 months, SD=2 years 5 months), seizures (n=12, onset mean=9 years 11 months, SD=3 years 7 months), and motor impairments (n=11, onset mean=9 years 10 months, SD=4 years 8 months).

Fourteen participants had in-person one-year follow up evaluations including Vineland and UBDRS by January 2020; three additional participants completed a remote interview, yielding additional Vineland data (Supplemental Table). The two participants with vision-

only phenotype were excluded from some analyses (leaving $n = 15$; Participants and Methods). VIQ data were obtained for 19 participants at baseline and eight participants at one-year follow-up. Descriptive statistics for the neuropsychological and disease severity scores are provided in Table 1. Mean verbal IQ and adaptive behavior scores and median UBDRS scores are not different between participants homozygous for the 1-kb deletion and those with other genotypes (excluding the 2 vision-only participants).

The cross-sectional relationships between the Vineland-3 ABC scores and VIQ and UBDRS scores are graphically depicted in Figure 1. Spearman correlations with 95% CI are shown in Table 2. Vineland-3 scores were strongly correlated with VIQ at both baseline ($n=19$; $r_{\text{range}}=[0.73 - 0.86]$; Figure 1A, Table 2) and follow-up ($n=8$; $r_{\text{range}}=[0.65 - 0.77]$; Table 2). However, reflecting a large mean difference in scores (at baseline, the sample average ABC score was 62.0 ± 26.8 , while the sample average VIQ was 78.6 ± 20.2), the concordance correlation coefficient was moderate ($\text{CCC} = 0.72$, $95\% \text{ CI}=[0.47, 0.86]$) and the 95% CI included values in the poor range (<0.50). A consistent result was obtained in the 1-year follow-up data, though the sample size was small: $\text{CCC} = 0.52$ $[-0.08, 0.84]$. Only one participant (at Time 1) scored at the floor of the VIQ test, so it seems unlikely that this is explained by the lower floor of the Vineland-3.

At baseline, Vineland-3 ABC scores were very strongly related with UBDRS scores ($n=22$). Correlations were negative for both the Physical ($r=-0.90$, $[-0.96, -0.77]$; Figure 1B) and CGI ($r=-0.87$, $[-0.94, -0.70]$; Figure 1D), and positive for the Capability ($r=0.93$, $[0.83, 0.97]$; Figure 1C) subdomains. All of these relationships were replicated in the subsample evaluated at 1-year follow-up ($n=17$; Table 2).

Vineland-3 norm-referenced (standard) scores were lower in older participants ($n=20$; Figure 2A). For each additional year of age, the expected difference in Vineland-3 score was $(-3.69 [-5.23, -2.16])$ points on the standard scale for the ABC, and similar differences across age were observed for the domains (Communication: $-4.13 [-6.10, -2.16]$, Daily Living Skills: $-4.49 [-5.84, -3.14]$, Socialization: $-3.09 [-5.20, -0.98]$). A consistent trend was observed within-participant in the subset with 1-year follow-up data ($n = 15$), such that average within-subject change over one year in ABC was $-3.66 [-7.13, -0.20]$ (Figure 2B, Supplemental Figure). The Communication domain (specifically the Written subdomain) and the Domestic and Interpersonal subdomains showed the most decrease at 1-year follow up. While the magnitude of these changes was within the range attributable to the measurement error of the instrument, the consistent decrease in normative scores across domains supports the interpretation that they are true declines.

Importantly, normative scores can decrease within subject because of skill loss, failure to gain skills at the age-expected rate, or both (Figure 2A,B). To discriminate amongst these, we evaluated the within-subject change in growth scale values (GSV) for each subdomain (Figures 2C,D). Average GSVs were stable for most sub-domains, reflecting equal proportions of participants with meaningful increases and decreases, indicating that decreases in norm-referenced scores were likely driven by both skill loss and slower-than-expected gains. Written Language declined between baseline and 1-year follow-up for

almost all participants, indicating that the decrease in norm-referenced score was likely due to loss of skills.

4 DISCUSSION

We evaluated the validity of the Vineland Adaptive Behavior Scales for use as an outcome measure in trials involving CLN3, using data from a prospective natural history study. At baseline, Vineland-3 Adaptive Behavior Composite (ABC) scores were quite variable, but were on average more than 2.5 standard deviations below the normative population mean, and more than one standard deviation below the values reported for a visual impairment reference sample in the Vineland-3 manual. This suggests that other factors related to CLN3 disease are impacting adaptive functioning. The validity of the Vineland-3 ABC was supported by strong and positive correlations with the established outcomes of verbal IQ and disease state as measured by the UBDRS. VIQ tended to be higher than ABC, supporting the idea that adaptive functioning is distinct from cognitive ability and is differentially affected by symptoms of the disorder. Consistent with the expected course of a neurodegenerative disease, older participants had lower Vineland-3 scores than younger participants, and small but widespread decreases were observed among the small sample of participants with 1-year follow-up data. Analysis of the person ability scores indicated that decreases over this relatively short follow-up period were caused by general plateauing rather than loss of skills, with the exception of a few specific subdomains (e.g., written) which may have been more directly affected by vision loss. Overall, these results support the application of the Vineland-3 adaptive behavior assessment (ABC) as an outcome measure in CLN3, but not its exchangeability with VIQ.

Therapeutic development for ultrarare pediatric disease faces scarcity of affected individuals for randomized controlled study designs and paucity of feasible and accessible outcome measures that account for development (*Assessing Neurocognitive Outcomes in Inborn Errors of Metabolism*, <https://www.fda.gov/media/96785/download>; Dickson et al., 2011; Shapiro et al., 2016). These challenges are compounded in pediatric neurodegenerative diseases, particularly those involving loss of functioning in multiple neurologic and sensory domains. One clinical outcome measure used and accepted by the U.S. Food and Drug Administration is cognitive ability assessment using standardized and validated methods such as IQ (Augustine et al., 2019; Krivitzky et al., 2009; Peters et al., 2004). Limitations such as floor effects relating to applicability in populations with significant intellectual disability and those with other concomitant disabilities, and in longitudinal assessment, are well recognized (Delaney et al., 2014; Dickson et al., 2011; Hessel et al., 2009; Janzen, Delaney, & Shapiro, 2017; Kronenberger, Harrington, & Yee, 2021; Shapiro et al., 2016). Adaptive behavior has therefore been proposed as an outcome measure as well and shown to be clinically relevant for therapeutic clinical trials in multiple conditions affecting neurocognitive status (Bolte & Poustka, 2002; Delaney et al., 2014; Freeman, Del'Homme, Guthrie, & Zhang, 1999; Hessel et al., 2009).

Previous assessment of adaptive behavior in children with CLN3 using the Scales of Independent Behavior-Revised (SIB-R) showed lower norm-referenced scores and floor effects for activities of daily living (Adams et al., 2006). Similar to the SIB-R, Vineland-3

is amenable to data collection from adult caregivers and includes norm references for infants to adults with broad ability range (Fletcher, 2014). In addition, the Vineland-3 offers the ability to collect data remotely, is more amenable to collection of data from non-English speaking participants (where interpreters were able to be used), and includes person ability scores, which directly indicate changes in ability rather than relative to age-based expectations (Farmer et al., 2020). This latter scoring option addresses the limitation of standardized scores to track stabilization or improvement in a pediatric neurodegenerative process (Farmer et al., 2020; Kronenberger et al., 2021). Results from this study showed that adaptive behavior scores obtained from the Vineland-3 should be considered as an outcome measure for neurodevelopmental functional assessment in CLN3, as its scores correlate strongly with UBDRS and VIQ scores. As shown in this study, however, adaptive behavior and VIQ are not interchangeable assessments.

At the individual level changes in Vineland-3 scores are heterogeneous at the one-year interval, suggesting that this duration is insufficient to capture significant progression as measured by these clinical parameters. Longitudinal trends of a slow decline in IQ, especially over one year, in CLN3 participants have been observed previously (Lamminranta et al., 2001). Small number of affected individuals and difficulty with longitudinal participation have limited calculations of annualized rate of changes for IQ in CLN3 to data from 5-year duration or cross-sectional cohorts (Kuper et al., 2018; Lamminranta et al., 2001). Comparisons of annualized rate of change for UBDRS scores derived from cross-sectional and longitudinal data suggested good agreement (Kwon et al., 2011). Further development of the UBDRS into a clinically meaningful disease staging tool proposed an average of 4-5 years separating the stages (Masten et al., 2020). Taken together, the findings from this and previous studies argue against CLN3 clinical trials as short as one to two years in duration due to limited changes in functioning in this natural history.

Limitations of this study include a bias of the enrollment of participants stable and healthy enough to travel and complete the visit. The trends in outcome measures reported here are based on a relatively small number of individuals with CLN3 disease. As our natural history study continues, additional longitudinal data and an increase in number of participants will provide insight into the reproducibility, applicability, and sensitivity of the reported findings.

In summary, data from this cohort show decreased adaptive behavior scores in individuals with CLN3, with meaningful concurrent correlations to verbal IQ and disease state. Thus, adaptive behavior is an applicable measure to be considered in future clinical trials. While this trend is also seen in a subset of participants at one-year follow-up, within participant changes were variable. The limited longitudinal data provide additional support for the Vineland-3 as continuing to track with disease severity. In addition, longitudinal changes in adaptive behavior show disease progression cannot be sufficiently measured over a one-year interval in cohorts similar to the one included in this study, a factor to be considered in the design of future therapeutic trials. It will be important to investigate how adaptive behavior, as well as other factors, change as children with CLN3 age and as the disease progresses.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENT

We dedicate the work described here to the participants, their family, and the family advocacy organizations (Beyond Batten Disease Foundation, Batten Disease Support and Research Association) for the motivation and inspiration they have provided. We thank Ms. Kisha Jenkins, RN (NICHD, NIH) for assisting in data collection and management; Dr. Jonathan W. Mink (University of Rochester Medical Center) for his expert inputs on UBDRS implementation. We thank NIH colleagues and staff who enabled the conduct of this study and the preparation of this manuscript. The NIH Intramural Research Program of NICHD (ZIA HD008989), NIMH (1ZICMH002961), the NIH Clinical Center, NINDS, and an NIH Clinical Center Bench-to-Bedside Award supported this work.

Funding

Division of Intramural Research, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (ZIA HD008989). Division of Intramural Research, National Institute of Mental Health (1ZICMH002961).

Data availability statement

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

REFERENCES

- Adams HR, de Bleeck EA, Mink JW, Marshall FJ, Kwon J, Dure L, ... Pearce DA (2006). Standardized assessment of behavior and adaptive living skills in juvenile neuronal ceroid lipofuscinosis. *Dev Med Child Neurol*, 48(4), 259–264. doi:10.1017/S0012162206000570 [PubMed: 16542512]
- Adams HR, Mink JW, & University of Rochester Batten Center Study, G. (2013). Neurobehavioral features and natural history of juvenile neuronal ceroid lipofuscinosis (Batten disease). *J Child Neurol*, 28(9), 1128–1136. doi:10.1177/0883073813494813 [PubMed: 24014508]
- Anderson GW, Goebel HH, & Simonati A (2013). Human pathology in NCL. *Biochim Biophys Acta*, 1832(11), 1807–1826. doi:10.1016/j.bbadis.2012.11.014 [PubMed: 23200925]
- Assessing Neurocognitive Outcomes in Inborn Errors of Metabolism. (<https://www.fda.gov/media/96785/download>). Paper presented at the Proceedings of Meeting held on April 16, 2015.
- Augustine EF, Beck CA, Adams HR, Defendorf S, Vierhile A, Timm D, ... Marshall FJ (2019). Short-Term Administration of Mycophenolate Is Well-Tolerated in CLN3 Disease (Juvenile Neuronal Ceroid Lipofuscinosis). *JIMD Rep*, 43, 117–124. doi:10.1007/8904_2018_113 [PubMed: 29923092]
- Bolte S, & Poustka F (2002). The relation between general cognitive level and adaptive behavior domains in individuals with autism with and without co-morbid mental retardation. *Child Psychiatry Hum Dev*, 33(2), 165–172. doi:10.1023/a:1020734325815 [PubMed: 12462353]
- Carrasco JLM, J. P. (2015). Concordance Correlation Coefficient for Repeated (and Non-Repeated) Measures. *R package*(version 1.2.1).
- Dang Do AN, Sinaii N, Masvekar RR, Baker EH, Thurm AE, Soldatos AG, ... Porter FD (2021). Neurofilament light chain levels correlate with clinical measures in CLN3 disease. *Genet Med*, 23(4), 751–757. doi:10.1038/s41436-020-01035-3 [PubMed: 33239751]
- Delaney KA, Rudser KR, Yund BD, Whitley CB, Haslett PA, & Shapiro EG (2014). Methods of neurodevelopmental assessment in children with neurodegenerative disease: Sanfilippo syndrome. *JIMD Rep*, 13, 129–137. doi:10.1007/8904_2013_269 [PubMed: 24190801]

- Dickson PI, Pariser AR, Groft SC, Ishihara RW, McNeil DE, Tagle D, ... Patterson MC (2011). Research challenges in central nervous system manifestations of inborn errors of metabolism. *Mol Genet Metab*, 102(3), 326–338. doi:10.1016/j.ymgme.2010.11.164 [PubMed: 21176882]
- Farmer C, Kaat A, Thurm A, Anselm I, Akshoomoff N, Bennett A, ... Miller J (2020). Person ability scores as an alternative to norm-referenced scores as outcome measures in studies of neurodevelopmental disorders. *AJIDD*. Retrieved from <http://aaid.org/publications/journals/articles-accepted-for-publication>
- Fletcher JM (2014). Alternative approaches to outcomes assessment: beyond psychometric tests. *Pediatr Blood Cancer*, 61(10), 1734–1738. doi:10.1002/pbc.24824 [PubMed: 24174385]
- Freeman BJ, Del'Homme M, Guthrie D, & Zhang F (1999). Vineland Adaptive Behavior Scale scores as a function of age and initial IQ in 210 autistic children. *J Autism Dev Disord*, 29(5), 379–384. doi:10.1023/a:1023078827457 [PubMed: 10587884]
- Freeman BJ, Ritvo ER, Yokota A, Childs J, & Pollard J (1988). WISC-R and Vineland Adaptive Behavior Scale scores in autistic children. *J Am Acad Child Adolesc Psychiatry*, 27(4), 428–429. doi:10.1097/00004583-198807000-00008 [PubMed: 3182599]
- Hayes S, & Farnill D (2003). Correlations for the Vineland Adaptive Behavior Scales with Kaufman Brief Intelligence Test in a forensic sample. *Psychol Rep*, 92(2), 573–580. doi:10.2466/pr0.2003.92.2.573 [PubMed: 12785645]
- Hessl D, Nguyen DV, Green C, Chavez A, Tassone F, Hagerman RJ, ... Hall S (2009). A solution to limitations of cognitive testing in children with intellectual disabilities: the case of fragile X syndrome. *J Neurodev Disord*, 1(1), 33–45. doi:10.1007/s11689-008-9001-8 [PubMed: 19865612]
- Janzen D, Delaney KA, & Shapiro EG (2017). Cognitive and adaptive measurement endpoints for clinical trials in mucopolysaccharidoses types I, II, and III: A review of the literature. *Mol Genet Metab*, 121(2), 57–69. doi:10.1016/j.ymgme.2017.05.005 [PubMed: 28506702]
- Johnson TB, Cain JT, White KA, Ramirez-Montealegre D, Pearce DA, & Weimer JM (2019). Therapeutic landscape for Batten disease: current treatments and future prospects. *Nat Rev Neurol*, 15(3), 161–178. doi:10.1038/s41582-019-0138-8 [PubMed: 30783219]
- Keith TZ, Fehrmann PG, Harrison PL, & Pottebaum SM (1987). The Relation Between Adaptive Behavior and Intelligence: Testing Alternative Explanations. *Journal of School Psychology*, 25, 31–43.
- Krivitzky L, Babikian T, Lee HS, Thomas NH, Burk-Paull KL, & Batshaw ML (2009). Intellectual, adaptive, and behavioral functioning in children with urea cycle disorders. *Pediatr Res*, 66(1), 96–101. doi:10.1203/PDR.0b013e3181a27a16 [PubMed: 19287347]
- Kronenberger WG, Harrington M, & Yee KS (2021). Projected Retained Ability Score (PRAS): A New Methodology for Quantifying Absolute Change in Norm-Based Psychological Test Scores Over Time. *Assessment*, 28(2), 367–379. doi:10.1177/1073191119872250 [PubMed: 31516019]
- Kuper WFE, van Alfen C, Rigterink RH, Fuchs SA, van Genderen MM, & van Hasselt PM (2018). Timing of cognitive decline in CLN3 disease. *J Inherit Metab Dis*, 41(2), 257–261. doi:10.1007/s10545-018-0143-x [PubMed: 29392585]
- Kuper WFE, van Alfen C, van Eck L, Huijgen BCH, Nieuwenhuis EES, van Brussel M, & van Hasselt PM (2019). Motor function impairment is an early sign of CLN3 disease. *Neurology*, 93(3), e293–e297. doi:10.1212/WNL.0000000000007773 [PubMed: 31182507]
- Kwon JM, Adams H, Rothberg PG, Augustine EF, Marshall FJ, Deblieck EA, ... Mink JW (2011). Quantifying physical decline in juvenile neuronal ceroid lipofuscinosis (Batten disease). *Neurology*, 77(20), 1801–1807. doi:10.1212/WNL.0b013e318237f649 [PubMed: 22013180]
- Lamminranta S, Aberg LE, Autti T, Moren R, Laine T, Kaukoranta J, & Santavuori P (2001). Neuropsychological test battery in the follow-up of patients with juvenile neuronal ceroid lipofuscinosis. *J Intellect Disabil Res*, 45(Pt 1), 8–17. doi:10.1046/j.1365-2788.2001.00288.x [PubMed: 11168772]
- Marshall FJ, de Blicke EA, Mink JW, Dure L, Adams H, Messing S, ... Pearce DA (2005). A clinical rating scale for Batten disease: reliable and relevant for clinical trials. *Neurology*, 65(2), 275–279. doi:10.1212/01.wnl.0000169019.41332.8a [PubMed: 16043799]

- Masten MC, Williams JD, Vermilion J, Adams HR, Vierhile A, Collins A, ... Mink JW (2020). The CLN3 Disease Staging System: A new tool for clinical research in Batten disease. *Neurology*, 94(23), e2436–e2440. doi:10.1212/WNL.0000000000009454 [PubMed: 32300063]
- Mole SE, & Cotman SL (2015). Genetics of the neuronal ceroid lipofuscinoses (Batten disease). *Biochim Biophys Acta*, 1852(10 Pt B), 2237–2241. doi:10.1016/j.bbadis.2015.05.011 [PubMed: 26026925]
- Morash VS, & McKerracher A (2017). Low reliability of sighted-normed verbal assessment scores when administered to children with visual impairments. *Psychol Assess*, 29(3), 343–348. doi:10.1037/pas0000341 [PubMed: 27183043]
- Peters C, Charnas LR, Tan Y, Ziegler RS, Shapiro EG, DeFor T, ... Krivit W (2004). Cerebral X-linked adrenoleukodystrophy: the international hematopoietic cell transplantation experience from 1982 to 1999. *Blood*, 104(3), 881–888. doi:10.1182/blood-2003-10-3402 [PubMed: 15073029]
- Preising MN, Abura M, Jager M, Wassill KH, & Lorenz B (2017). Ocular morphology and function in juvenile neuronal ceroid lipofuscinosis (CLN3) in the first decade of life. *Ophthalmic Genet*, 38(3), 252–259. doi:10.1080/13816810.2016.1210651 [PubMed: 27486012]
- R-Core-Team. (2020). R: A language and environment for statistical computing. Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Shapiro E, Bernstein J, Adams HR, Barbier AJ, Buracchio T, Como P, ... Mulberg AE (2016). Neurocognitive clinical outcome assessments for inborn errors of metabolism and other rare conditions. *Mol Genet Metab*, 118(2), 65–69. doi:10.1016/j.ymgme.2016.04.006 [PubMed: 27132782]
- Signorell A, & al., e. (2020). DescTools: Tools for Descriptive Statistics (Version 0.99.36). Retrieved from <https://cran.r-project.org/package=DescTools>
- Sparrow S, Cicchetti D, & Saulnier C (2016). Vineland Adaptive Behavior Scales—Third Edition (Vineland-3). In. San Antonio, TX: Pearson.
- Specchio N, Ferretti A, Trivisano M, Pietrafusa N, Pepi C, Calabrese C, ... Vigevano F (2020). Neuronal Ceroid Lipofuscinosis: Potential for Targeted Therapy. *Drugs*. doi:10.1007/s40265-020-01440-7
- Wasserstein R, Schirm A, & Lazar N (2019). Moving to a World Beyond “ $p < 0.05$ ”. *The American Statistician*, 73, 1–19. doi:10.1080/00031305.2019.1583913
- Wechsler D (2008). WAIS-IV administration and scoring manual. In. San Antonio, TX: Psychological Corporation.
- Wechsler D (2014). Wechsler intelligence scale for children—Fifth Edition (WISC-V). In. Bloomington, MN: Pearson.

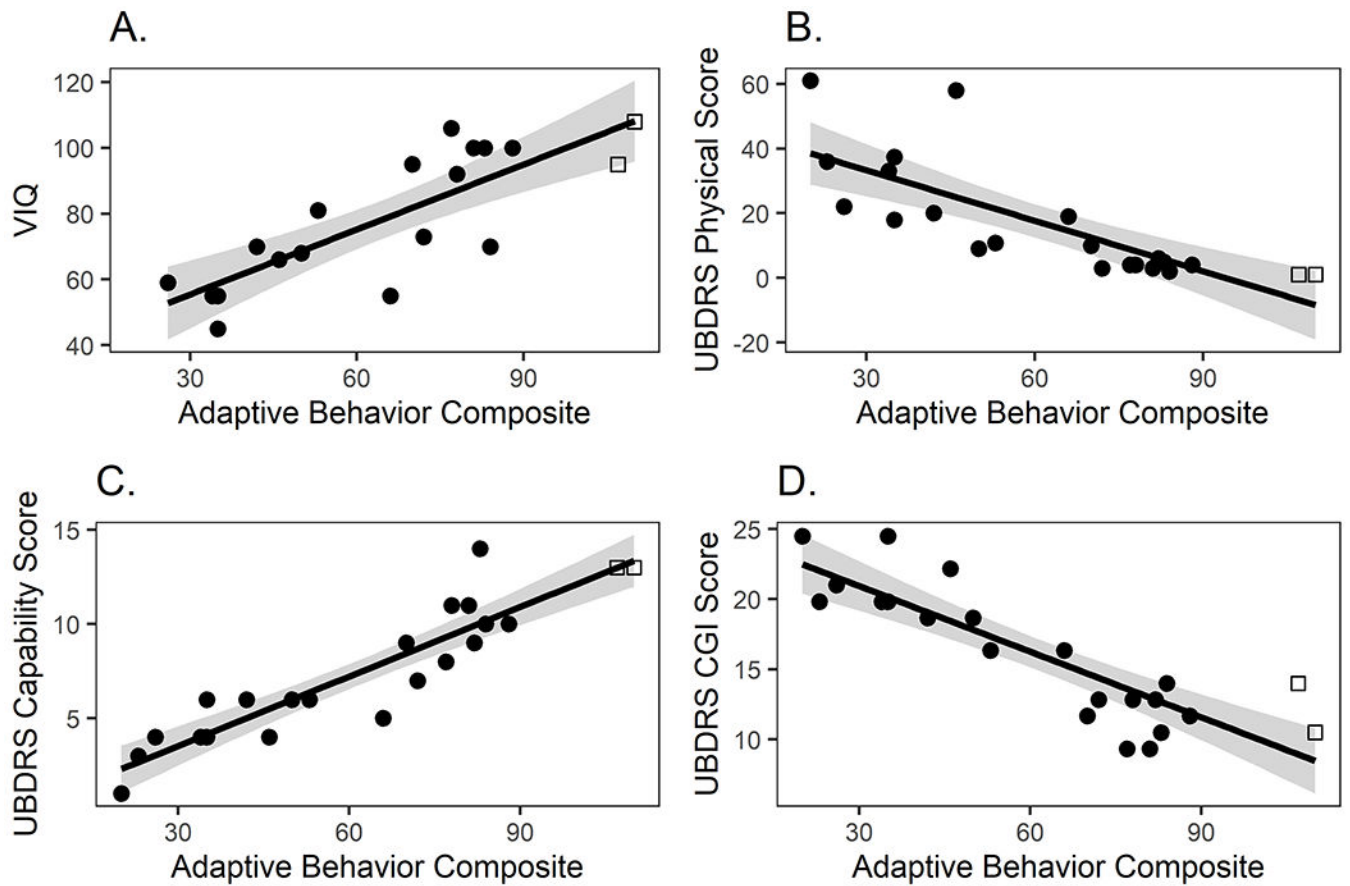


Figure 1.

Relationships of the Vineland-3 Adaptive Behavior Composite scores in CLN3 disease.

ABC scores strongly correlate with established measures, VIQ (A) and UBDRS sub-domain scores (B-D) and. VIQ is a standard score; UBDRS scores are weighted.

Circles: Participants with classic CLN3 phenotype. Squares: Participants with vision-only phenotype. Black line: linear regression. Shaded gray area: 95% CI. ABC: adaptive behavior composite. CGI: clinical global impression. UBDRS: Unified Batten Disease Rating Scale. VIQ: verbal intelligence quotient.

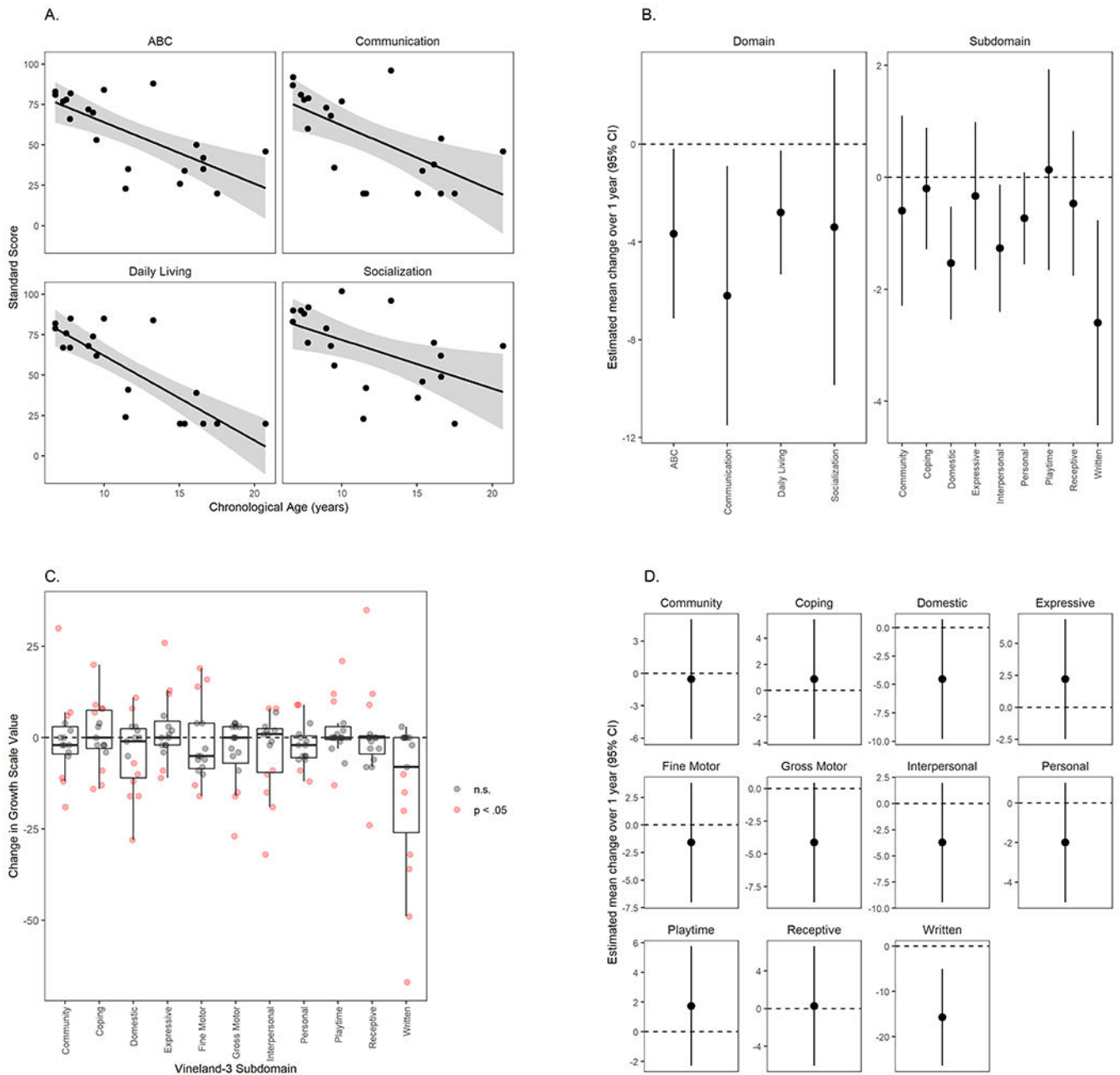


Figure 2. Adaptive behavior scores at baseline and 1-year follow-up in individuals with CLN3-related disorders. **A.** Cross-sectional relationship (baseline only) of Vineland-3 adaptive behavior composite and domain scores to age (n=20, excluding vision-only participants). Line: fitted curve for CLN3. Circles: individual data points. Shaded area: 95% CI. **B.** Model-estimated mean within-subject change over 1-year follow-up in Vineland-3 domain (standard) and subdomain (V-scale) scores (n=15, excluding vision-only participants). ABC: adaptive behavior composite. **C.** Within-subject change over 1-year follow-up in Vineland-3 growth scale values in CLN3 disease (n=15, excluding vision-only participants). Change scores were compared to the manual-based average SEM for each subdomain, and values that

exceeded $2 \times \text{SEM}$ are shaded pink. **D.** Model-estimated mean within-subject change over 1-year follow-up in Vineland-3 growth scale values in CLN3 disease (n=15, excluding vision-only participants). Score of zero represents no change in ability; negative scores indicate a decrease in ability.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Neuropsychological and disease severity descriptive statistics for CLN3 cohort.

	All			Homozygous 1-kb deletion			Other genotypes ^a			
	N	Range	Mean (SD) or Median [IQR]	N	Range	Mean (SD) or Median [IQR]	N	Range	Mean (SD) or Median [IQR]	
Baseline										
Vineland-3	<i>ABC</i>	22	20-110	61.9 (26.8)	10	26-82	55.1 (21.2)	10	20-88	59.4 (26.1)
	<i>Comm</i>	22	20-113	59.6 (30.0)	10	20-87	51.4 (25.2)	10	20-96	58.5 (30.2)
	<i>DLS</i>	22	20-106	57.4 (29.9)	10	20-85	49.8 (28.6)	10	20-85	55.5 (26.6)
	<i>Social</i>	22	20-120	70.6 (26.7)	10	36-92	65 (18.8)	10	20-102	68 (29.9)
VIQ		19	45-108	78.6 (20.2)	9	55-100	75 (17.6)	8	45-106	76.9 (22.7)
	<i>Physical</i>	22	1-61	9.5 [4;21.5]	10	3-33	10.4 [6.8;19.5]	10	2-61	12 [4;37]
UBDRS	<i>Capability</i>	22	1-14	6.5 [4.3;10]	10	4-11	6 [4.5;9]	10	1-14	6.5 [4.3;9.5]
	<i>CGI</i>	22	9.3-24.5	15.2 [12.0;19.8]	10	9.3-21	17.5 [12.8;19.5]	10	9.3-24.5	15.2 [12.0;21.6]
1-year Follow-up										
Vineland	<i>ABC</i>	17	24-117	64 (26.5)	8	33-84	56.6 (20.2)	7	24-76	58.9 (22.9)
	<i>Comm</i>	17	20-109	58.9 (27.4)	8	20-83	51.6 (22.0)	7	20-79	53.9 (24.5)
	<i>DLS</i>	17	20-120	62.3 (29.4)	8	20-89	52.4 (28.2)	7	31-79	59.1 (18.9)
	<i>Social</i>	17	20-110	71.4 (27.1)	8	34-94	66.9 (22.05)	7	20-90	65.7 (29.2)
VIQ		8	72-103	89.6 (11.9)	3	72-92	80 (10.6)	3	81-103	93 (11.1)
	<i>Physical</i>	13	1-36	7 [6;13]	7	3-36	9 [5.5;28.5]	5	6-13	7 [6;8]
UBDRS	<i>Capability</i>	14	4-14	7 [6;7.8]	7	4-8	7 [5.5;7]	5	5-9	7 [6;7]
	<i>CGI</i>	14	11-22	15.5 [14;16.8]	7	15-22	16 [15.5;18.5]	5	12-17	14 [14;16]

^aExcluding 2 participants with vision-only phenotype.

ABC: adaptive behavior composite. CGI: clinical global impression. Comm: communication. DLS: daily living skills. IQR: interquartile range. SD: standard deviation. UBDRS: Unified Batten Disease Rating Scale. VIQ: verbal intelligence quotient.

Spearman Rho [95% CI] for relationship between Vineland-3 and established outcome (VIQ or UBDRS) at baseline and one-year follow-up.

Table 2.

	VIQ	UBDRS		
		Physical	Capability	CGI
Baseline				
ABC	0.81 [0.57; 0.92]	-0.90 [-0.96; -0.77]	0.93 [0.83; 0.97]	-0.87 [-0.94; -0.71]
Communication	0.86 [0.66; 0.94]	-0.85 [-0.93; -0.66]	0.91 [0.78; 0.96]	-0.90 [-0.96; -0.78]
Daily Living Skills	0.73 [0.41; 0.89]	-0.86 [-0.94; -0.69]	0.89 [0.76; 0.95]	-0.84 [-0.93; -0.64]
Socialization	0.74 [0.42; 0.89]	-0.89 [-0.95; -0.76]	0.86 [0.68; 0.94]	-0.82 [-0.92; -0.61]
1-year Follow-Up				
ABC	0.72 [0.04; 0.95]	-0.79 [-0.93; -0.42]	0.81 [0.48; 0.94]	-0.75 [-0.91; -0.36]
Communication	0.65 [-0.1; 0.93]	-0.72 [-0.91; -0.29]	0.74 [0.35; 0.91]	-0.66 [-0.88; -0.20]
Daily Living Skills	0.76 [0.12; 0.95]	-0.77 [-0.93; -0.39]	0.83 [0.54; 0.94]	-0.67 [-0.89; -0.22]
Socialization	0.77 [0.15; 0.96]	-0.80 [-0.94; -0.45]	0.79 [0.44; 0.93]	-0.77 [-0.92; -0.41]

ABC = Adaptive Behavior Composite. CGI = Clinician Global Impression. UBDRS = Unified Batten Disease Rating Scale. VIQ = Verbal Intelligence Quotient. Individuals with missing data were excluded on a pairwise basis, sample sizes were as follows: Vineland with IQ, n = 19 (baseline) and n = 8 (1-year follow-up). Vineland with UBDRS, n = 22 (baseline) and n = 17 (1-year follow-up).