American Journal on Intellectual and Developmental Disabilities

Characterizing developmental and behavioral profiles in developmental synaptopathies to inform clinical trial endpoints

--Manuscript Draft--

Manuscript Number:	AJIDD-D-25-00003R1
Article Type:	Invitation Only - Developmental Synaptopathies
Keywords:	tuberous sclerosis complex (TSC); PTEN Hamartoma Tumor Syndrome (PHTS); Phelan McDermid Syndrome (PMS); neuropsychological assessment; clinical trial readiness
Corresponding Author:	Cristan A. Farmer National Institute of Mental Health Bethesda, MD UNITED STATES
First Author:	Latha Valluripalli Soorya
Order of Authors:	Latha Valluripalli Soorya
	Camille W Brune
	Cristan A. Farmer
	Edith V. Ocampo
	Natalie I Berger
	Deborah A Pearson
	Robyn M Busch
	Patricia Klaas
	Paige Siper
	Kristn Currans
	Amanda C Gulsrud
	Jennifer M Phillips
	Rajna Filip-Dhima
	Sarah E O'Kelley
	Thomas W Frazier
	Tess Levy
	Allison L Wainer
	Joseph D Buxbaum
	Craig M Powell
	Jonathan A Bernstein
	Simon K Warfield
	Darcy A Krueger
	E. Martina Bebin
	Hope Northrup
	Shafali S Jeste
	Alexander Kolevzon
	Elizabeth Berry-Kravis
	Mustafa Sahin

	Siddharth Srivastava
	Audrey Thurm
Manuscript Region of Origin:	UNITED STATES
Abstract:	Background The Developmental Synaptopathies Consortium is a multi-site natural history network studying rare, neurogenetic syndromes associated with synaptic dysfunction and developmental delays. One aim of the Consortium is clinical trial readiness, including identifying clinical concepts and validating their measurement. Methods We evaluated the scope and limitations of conventional cognitive and behavioral measurement strategies in 2–21-year-olds with Phelan-McDermid syndrome (PMS, N=98), Tuberous Sclerosis Complex (TSC, N=98) and PTEN Hamartoma Tumor syndrome (PHTS, N=69). Results On average, intellectual disability (ID) severity was severe-to-profound in PMS, mild-to-moderate for TSC, and borderline (or absent) in PHTS. Severity of ID invalidated the use of many assessments, including standardized autism diagnostic measures. Conclusions These results will inform trial planning for these and other similarly medically complex neurodevelopmental conditions.

Abstract

Background: The Developmental Synaptopathies Consortium is a multi-site natural history network studying rare, neurogenetic syndromes associated with synaptic dysfunction and developmental delays. One aim of the Consortium is clinical trial readiness, including identifying clinical concepts and validating their measurement. **Methods:** We evaluated the scope and limitations of conventional cognitive and behavioral measurement strategies in 2–21-year-olds with Phelan-McDermid syndrome (PMS, N=98), Tuberous Sclerosis Complex (TSC, N=98) and PTEN Hamartoma Tumor syndrome (PHTS, N=69). **Results:** On average, intellectual disability (ID) severity was severe-to-profound in PMS, mild-to-moderate for TSC, and borderline (or absent) in PHTS. Severity of ID invalidated the use of many assessments, including standardized autism diagnostic measures. **Conclusions:** These results will inform trial planning for these and other similarly medically complex neurodevelopmental conditions.

Keywords: autism spectrum disorder (ASD), tuberous sclerosis complex (TSC), PTEN Hamartoma Tumor Syndrome (PHTS), Phelan McDermid Syndrome (PMS), intellectual disability (ID), neuropsychological assessment, neurodevelopmental disorders, clinical trial readiness

Introduction

Advances in genomics have pointed to considerable overlap in genetic risk factors for neurodevelopmental disorders such as intellectual disability (ID) and autism spectrum disorder (ASD) (Sanders et al., 2019). These advances present a promising opportunity to interrogate shared molecular and disease mechanisms informing medical, physiological, psychiatric, cognitive, and behavioral outcomes, which may reveal common treatment pathways for related conditions.

The Developmental Synaptopathies Consortium (DSC) was formed and funded by the Rare Disease Clinical Research Network (RDCRN) in 2014 to advance this genetics-first approach to treatment development. The DSC studied three genetic syndromes associated with genes governing early synaptogenesis (Dölen & Sahin, 2016) and associated with disruptions or dysregulation of the mTOR pathway (Sahin & Sur, 2015; Winden et al., 2018): 1) Phelan-McDermid Syndrome (PMS), 2) Tuberous Sclerosis Complex (TSC), and 3) PTEN Hamartoma Tumor Syndrome (PHTS) (Sahin & Sur, 2015; Winden et al., 2018). Mutations in *SHANK3* impact pathways (e.g., mTOR signaling) common to multiple monogenic forms of ID and ASD, including tuberous sclerosis (TSC) and PTEN-related disorders (PRD) (Darnell et al., 2011; Sakai et al., 2011) possibly by affecting signals from the glutamate receptors (Nisar et al., 2022).

Phelan McDermid Syndrome (PMS)

Individuals with PMS are missing a copy (i.e., haploinsufficient) of *SHANK3* (Anderlid et al., 2002; Wilson et al., 2003), due to either a pathogenic *SHANK3* sequence variants or a deletion of 22q13 containing *SHANK3*. *SHANK3* is a critical scaffolding protein responsible for promoting the growth and maturation of dendritic spines (Phelan & McDermid, 2012; Sala et al., 2001;

Sarasua et al., 2011). Prevalence rates of PMS are not well established, with estimates ranging between 1 in 20,000 and 1 in 50,000 in the general population (Betancur & Buxbaum, 2013; Phelan & McDermid, 2012). Physical features may include non-specific facial dysmorphology, joint hypermobility, and dysplastic fingernails and toenails, but are largely variable except for early presence of hypotonia (Schön et al., 2023). Gastrointestinal issues are prevalent, and about a third of individuals with PMS have seizures at some time in life (Levy, Gluckman, et al., 2024). PMS is often associated with pervasive delays in all areas of development, especially in early communication and motor skills (Dille et al., 2022; Farmer et al., 2024; Kohlenberg et al., 2020; Kolevzon et al., 2019; Schön et al., 2023; Soorya et al., 2013).

Moderate-to-profound ID is common among individuals with PMS (Srivastava et al., 2023) with estimates around 77% (Scalisi et al., 2022; Soorya et al., 2013). Some studies have reported decreasing cognitive ability with age, but more recent work calls into question whether this is attributable to a true loss of skills, slower-than-expected acquisition of skills, and/or floor effects of measures (Soorya et al., 2018; Srivastava et al., 2023). ASD prevalence estimates range 50–80% (Oberman et al., 2015; Xu et al., 2020), though they are confounded by the high rate of moderate-to-profound ID (Soorya et al., 2018).

Regression of previously acquired skills (e.g., communication, motor) in PMS may occur through adulthood (Dille et al., 2022; Kohlenberg et al., 2020). Rates of regression vary, but loss of at least one developmental skill is reported in over a third of individuals with PMS (Dille et al., 2022; Soorya et al., 2013; Farmer et al., in press). Adolescent and adult neuropsychiatric regression has also been observed among individuals with PMS, with an onset often emerging between 15 and 20 years of age (Kohlenberg et al., 2020; Kolevzon et al., 2019). Several

consensus guidelines and recommendations for management of PMS have been published (Srivastava et al., 2023; van Eeghen et al., 2023; van Ravenswaaij-Arts et al., 2023).

Tuberous Sclerosis Complex (TSC)

TSC is an autosomal dominant disorder caused by pathogenic sequence variants in the *TSC1* or *TSC2* genes located on chromosomes 9q34 and 16p13.3, respectively (Crino et al., 2006). *TSC1* and *TSC2* play a critical role in cell growth and division; mutations in these genes result in a multi-system disorder characterized by noncancerous tumors in the brain, heart, kidneys, lungs, and skin (Curatolo et al., 2008). The prevalence of TSC is estimated at 1 in 6,000 to 1 in 10,000 (O'Callaghan et al., 1998). TSC is typically identified because of seizures, which occur in 80 to 90% of individuals (Krueger et al., 2013; Levine et al., 2023), or kidney and heart hamartomas, though both skin lesions (e.g., shagreen patches) and macrocephaly are also commonly observed in infants with TSC (Levine et al., 2023). TSC is associated with several clusters of symptoms characterized as autism-like, dysregulated behavior, eat/sleep, mood/anxiety, neuropsychological, overactive/impulsive, and scholastic (de Vries et al., 2023).

Prevalence estimates for ID in TSC range from 44 to 64% (Goh et al., 2005; Jansen et al., 2008; Joinson et al., 2003), and at least one early epidemiological study indicated that the IQ distribution had a strong left skew, with 31% of the sample falling into the profound range (Joinson et al., 2003). However, studies with later-born cohorts have documented more normal distributions of IQ despite more severe seizure activity, suggesting an important role for the early detection and treatment of epileptic spasms in reducing disability associated with TSC (Tye et al., 2018). One meta-analysis documented the rate of ASD diagnosis estimates ranging 20–69% with increased prevalence of autism associated with seizure onset before age 2 (Specchio et al., 2020). Updated recommendations for TSC include early detection of tumors in the brain, routine

skin exams, gathering baseline EEG activity, monitoring for neurological, psychiatric, and behavioral symptoms to begin early intervention (Northrup et al., 2021), medication for early seizure management (de Saint Martin et al., 2022; Northrup et al., 2021; Yum et al., 2013), and annual screening for TSC-Associated Neuropsychiatric Disorders (TAND) (de Vries et al., 2023).

PTEN Hamartoma Tumor Syndrome (PHTS)

PHTS is caused by pathogenic sequence variants in *PTEN*, a ubiquitous tumor suppressor gene located on chromosome 10 (Hobert & Eng, 2009). *PTEN* codes for a protein that helps regulate a range of cellular functions including cell growth and division, apoptosis, and cell migration (Maehama et al., 2001; Maehama & Dixon, 1999). The prevalence of PHTS is estimated at approximately 1 in 200,000 to 1 in 250,000 (Hobert & Eng, 2009). PHTS is typically identified due to the macrocephaly that occurs in nearly all individuals with this condition (National Organization of Rare Disorders, 2023). Medical complications from PHTS include an increased risk for benign hamartomas as well as certain cancers including thyroid, colon, and breast cancer with specific, lifespan screening recommendations (Dhawan et al., 2025). The neuropsychological profile of PHTS is heterogeneous. ID prevalence estimates range from 4 to 32% in the mild-to-moderate range (Busch et al., 2013; Hansen-Kiss et al., 2017) and neuropsychological deficits in attention, working memory, and executive functioning are commonly reported in PHTS (Butler et al., 2005; Frazier et al., 2015; Varga et al., 2009). A large minority (between 8 and 25%) of children with PHTS are diagnosed with ASD (Butler et al., 2005; Ciaccio et al., 2019; Cummings et al., 2022; Varga et al., 2009), with a symptom profile characterized by more sensory issues and less severe social communication symptoms than idiopathic ASD (Busch et al., 2019).

The Developmental Synaptopathies Consortium (DSC)-I

A primary goal of the DSC was clinical trial readiness, for which natural history research is an essential initial step (Food and Drug Administration, 2019). Natural history studies are used to define the patient population and identify clinically meaningful concepts of interest, i.e., aspects of the condition that affect how a patient feels, functions, or survives (Food and Drug Administration, 2022). Natural history studies are also used to identify, refine, or develop clinical outcome assessments and biomarkers that correlate with those concepts of interest. Finally, natural history studies may even provide external control data for rare, complex conditions where controlled treatment trials are infeasible.

The goal of the current analysis is to inform the identification of clinically meaningful concepts and approaches to feasible and valid measurement of cognitive and behavioral phenotypes associated with these developmental synaptopathies. To accomplish this, we describe the results from baseline neurodevelopmental phenotyping performed in these natural history studies, with a focus on the feasibility and validity of assessments and their resulting data. This report is complementary to other work describing the individual disease phenotypes themselves (Busch et al., 2019, 2023; Farach et al., 2019; Levy et al., 2022; Levy, Farmer, et al., 2024).

Methods

Participants

Participants were recruited from six U.S. sites for Phelan-McDermid syndrome (Levy et al., 2022), five sites for TSC (Farach et al., 2019), and four sites for PHTS (Busch et al., 2019). The inclusion criteria for all studies were English speaking individuals with confirmed pathogenic findings of 1) PMS: pathogenic/likely pathogenic *SHANK3* sequence variant or deletion of 22q13 that affects the *SHANK3* gene, 2) TSC: *TSC1* or *TSC2* pathogenic/likely

pathogenic variants, or 3) PHTS: pathogenic/likely pathogenic variants in *PTEN* or deletions of *PTEN*. Participants were ages 3–21 years, except for two participants with TSC (24- and 35-months-old). Parents/guardians provided consent, and when appropriate participants provided assent. The protocols were approved by a central Institutional Review Board.

Study Procedures

Each protocol enacted a systematic neurodevelopmental assessment framework designed for investigating rare genetic conditions associated with ID (Soorya et al., 2018). This framework employs multiple methods (observer-rated outcomes, clinician-rated outcomes, and direct assessment) and multiple assessments (i.e., different measures of the same concept) to yield estimates of functioning across the full range of phenotypic expression observed in these heterogeneous populations. Based on the extant literature, the investigators' clinical experience with the conditions, and input from parent advocacy group members who served on the steering committees, clinical outcome assessments for the following concept domains were selected: intellectual functioning (including processing speed, working memory), visual motor integration, and executive functioning, language, adaptive functioning, challenging behavior, and ASD symptoms. Supplementary materials detailing the assessment framework and the original list of measures may be found online at https://doi.org/10.31234/osf.io/gqvr8_v1.

Results presented herein are data collected at the first (baseline) of three annual visits in this prospective, natural history study. Psychological and behavioral evaluations were conducted at each site by licensed clinicians or psychometrists (technicians supervised by a licensed psychologist). Annual network-wide consensus coding meetings were held to review reliability and measurement challenges. Participants typically completed in-person assessments over 1–2 days while parents/caregivers completed questionnaires and interviews.

Measures

The measures comprising the systematic neurodevelopmental assessment framework are described below by domain. The selection of appropriate clinical outcome assessments and endpoints (i.e., the scores used from those assessments) required the consideration of several factors, including: 1) the standardization age range (see Supplementary Materials at https://doi.org/10.31234/osf.io/gqvr8_v1), 2) the impact of planned modifications to standardized procedures on scoring and interpretation, and 3) the minimum developmental levels required to evaluate each developmental domain (e.g., social communication skills in children with developmental levels >12 months) (Soorya et al., 2018).

Intellectual functioning, language, and adaptive behavior

Intellectual functioning. Direct assessments of intellectual functioning were selected from the following hierarchy based on clinical judgment of the participant's chronological age, language level, and estimated developmental level:

1) Stanford Binet, 5th edition (Roid & Pomplun, 2012). The Stanford Binet is a traditional IQ test appropriate for individuals aged 2–85+ years. The Stanford-Binet offers Full Scale IQ (FSIQ), Nonverbal IQ (NVIQ), and Verbal IQ (VIQ) standard scores which range 40 to 160 (population mean±SD=100±15).
2) Differential Ability Scales, 2nd edition (DAS-II) (Elliott, 2007). The DAS-II content is transitionary between developmental and traditional IQ tests, depending on the form (Early Years or School Age). The DAS-II is normed for ages 2 years, 6 months to 17 years, 11 months. The score used for FSIQ on the DAS-II was General Conceptual Ability Core Cluster; NVIQ and VIQ were the Nonverbal and Verbal Composite Core

Clusters, respectively. These standard scores range from 45–165 for FSIQ and 50–150 for NVIQ and VIQ (population mean±SD=100±15).

3) Mullen Scales of Early Learning (MSEL) (Mullen, 1995). The MSEL is a developmental test normed for children birth to 5 years, 8 months and was used to provide estimates of cognitive functioning when the other measures were not feasible. To facilitate comparability with FSIQ, NVIQ, and VIQ from the other tests, developmental quotients (DQ; also known as ratio IQs) were calculated using the average age equivalents for the Fine Motor and Visual Reception (nonverbal; NVDQ) or Expressive Language and Receptive Language (verbal; VDQ) subscales (FSDQ = average of NVDQ and VDQ). DQs have no population distribution and a natural floor of 0, with a ceiling of infinity (see Farmer, Thurm, et al., in press).

For some participants with significant ID, it was not possible to achieve a basal score on the age-appropriate test, in which case the next-lower test in the hierarchy was administered. This resulted in some out-of-age-range testing with the MSEL. This commonly used approach has been shown to yield scores that correlate with but exhibit varying levels of mean difference with IQ scores (Bishop et al., 2011; Farmer et al., 2016). For all cognitive scores, higher values indicate better relative performance. For descriptive purposes, scores are also categorized by the commonly used ID designations (American Medical Association, 2018): No ID (standard score≥70), Mild (50 - 69), Moderate (35 - 49), Severe/Profound (below 35/under 20).

Language. Direct assessments of receptive and expressive language skills were administered to participants with developmental and learning readiness skills (e.g., attending to and discriminating 2D stimuli). The Peabody Picture Vocabulary Test-3 (PPVT-4) (Dunn & Dunn, 2007) and Expressive Vocabulary Test-2 (EVT-2) (Williams, 2007) were used to directly

assess receptive and expressive vocabulary, respectively. These tests are normed for ages 2 years, 6 months to 90+ years, and each provides a standard score ranging 20–160 (population distribution 100±15). Higher scores indicate better relative performance.

Adaptive Behavior. The Vineland Adaptive Behavior Scales, second edition (Vineland-II) (Sparrow et al., 2005) was used to evaluate adaptive behavior. For consistency with previous studies, the parent/caregiver survey form was used in TSC and PHTS, and the semi-structured clinician-led interview form was used for PMS. Both forms provide domain-level standard scores that range from 20–160 (population distribution 100±15) and subdomain scaled scores ranging 1–23 (population distribution 15±3). Higher scores indicate better relative adaptive functioning. Due to the prominence of motor delays in the PMS phenotype (Frank, 2021), the Motor Skills domain was administered for PMS participants regardless of participant age (normative data are available only through age 6 years 11 months and older adults). As a result, Motor Skills age equivalents, but not standard scores, were available for these individuals.

Behavioral measures

Challenging behaviors. The validity of clinical outcome assessments for challenging behaviors is compromised for individuals with intellectual and developmental disabilities (IDD) because diagnostic criteria are primarily defined in the context of typical development. For instance, behavior such as running in inappropriate situations may reflect alternative (maladaptive) communication patterns in younger and/or cognitively impaired populations but may be captured as overactivity by standardized challenging behavior measures. With this limitation in mind, we selected two widely used caregiver-report forms of challenging behaviors:

the Aberrant Behavior Checklist-Community Version (ABC-C) (Aman & Singh, 1994) and the Child Behavior Checklist/Adult Behavior Checklist (CBCL/ABCL) (Achenbach, 1999).

The ABC-C was originally developed for individuals with IDD in residential settings, having since been revised for community settings, with psychometric evaluations in a variety of samples with all levels of ID and other developmental disabilities such as ASD (Aman & Singh, 2017). The ABC-C yields sum total scores on five subscales: Irritability, Social Withdrawal, Stereotypic Behavior, Hyperactivity/Noncompliance, and Inappropriate Speech. The subscale sum scores are typically used for research purposes, though some comparison data are provided by the test developer for the calculation of norm-referenced scores. In this study, we used the ABC community samples (Marshburn & Aman, 1992) which included some participants with mild ID to guide interpretation of the clinical significance of ABC subscale scores.

The CBCL/ABCL is a screening tool for emotional and behavioral problems, with forms standardized for ages 1.5 years through adulthood. Sex-based normative data for each age-based form are provided, based on general population development samples. The CBCL yields T-scores with a range of 50–100 (population distribution 50±10). Higher scores indicate more symptoms relative to age peers. The CBCL/ABCL was standardized in typically developing populations with studies pointing to concerns with reliability and validity in those with moderate, severe, and profound ID (Koskentausta et al., 2004) and distinct patterns observed in children with ASD and NVIQ in the ID range (Piergies et al., 2022). Despite these concerns about the validity of the instrument, we included the CBCL to enable comparisons with extant literature, particularly in PHTS and TSC (Dekker et al., 2002; Dovgan et al., 2019; Esbensen et al., 2018; Pandolfi et al., 2009). However, to improve the assumption of measurement invariance across the development sample and the current studies, thereby reducing the potential to over-or under-

estimate behavioral challenges in children with severe to profound ID, we considered CBCL administrations to be valid for analysis only when the child had an NVIQ≥70.

ASD. ASD phenomenology was measured through multiple methods including caregiver report, developmental history, and direct, clinician-administered assessments. Within the PMS and TSC studies, participants received all clinician- and caregiver-rated instruments described below. In the PHTS study**,** for which groups were categorized by ASD status for an associated clinical trial (NCT#02991807), clinician-administered instruments were conducted after a DSM-5 checklist interview and administered only for the subset of participants who presented with suspected autism or had a prior ASD diagnosis (PHTS/ASD) (Busch et al., 2019).

The Social Responsiveness Scale (SRS-2) (Constantino & Gruber, 2005) is a caregiver-reported ASD screening instrument that is frequently used in neurodevelopmental research (Busch et al., 2023). There are three forms (preschool, school age, adult) that are administered based on chronological age. While T-scores are available for several subscales on the SRS-2, it was not validated in individuals with ID and/or significant expressive language delays (Constantino & Gruber, 2005; Hus et al., 2013) and research has demonstrated reduced validity in ID (Havdahl et al., 2016), including one study in PMS (Gergoudis et al., 2020) and another in fragile X syndrome (Kidd et al., 2020). To address these limitations, this study utilizes a 16-item short form that retains only items that exhibited measurement invariance across sex, age, expressive language, adaptive behavior, and NVIQ (Lyall et al., 2021; Sturm et al., 2017). While a limitation of the short form is a lack of clinical cut-off scores, we felt that the ability to use the short form score as an index of severity made it the preferable metric. Additionally, previous work demonstrated the validity of this approach for the PMS study (Gergoudis et al., 2020). Raw sum scores on the short form range from 0 to 48, where higher scores indicate more symptoms.

The Repetitive Behavior Scale-Revised (RBS-R) is a caregiver-report questionnaire developed for individuals with ID with and without ASD to characterize repetitive behaviors in IDDs (including ASD) (Bodfish et al., 2000). The RBS-R was originally developed with six subscales: Ritualistic Behavior, Sameness Behavior, Self-Injurious Behavior, Stereotyped Behavior, Compulsive Behavior and Restricted Behavior; though current standards, used here, condense the Ritualistic Behavior and Sameness Behavior subscales into one (Rituals/Sameness) (Lam & Aman, 2007). Higher raw sum scores indicate more severe problems.

The Autism Diagnostic Observation Schedule (2nd edition) (ADOS-2) is a clinician-administered assessment of symptoms related to an ASD diagnosis. The appropriate ADOS-2 module is selected based on the participant's age and language level. The Autism Diagnostic Interview, Revised (ADI-R) is a clinician-rated semi-structured interview about past and current symptoms related to an ASD diagnosis. Both the ADOS-2 and ADI-R yield sum scores in several domains, which are converted to cutoffs reflecting likelihood of an ASD diagnosis. The ADOS-2 and ADI-R were both administered by research-reliable clinicians, who underwent periodic reliability checks during the study. The DSM-5 checklist was completed by licensed clinicians (e.g., neurologist, psychiatrist, psychologist) indicating their judgment of the presence or absence of each of the ASD diagnostic criteria, based on all available information. Together, these tools comprise the gold-standard research assessment for ASD.

Of note, the development samples for the ADOS-2 and ADI-R were largely without severe-to-profound ID, and both the test developers (Lord et al., 2012; Rutter et al., 2003) and other researchers (Risi et al., 2006; Thurm et al., 2019) have noted that the specificity of the instruments can be much worse for individuals with low mental age. For example, the ADI-R and ADOS-2 are standardized for individuals with a mental age ≥18 months, but improved

specificity is found for mental age ≥2 years (Kim et al., 2013; Lord et al., 1993). The ADOS-2 has an additional criterion of independent walking. We followed manual-based recommendations, requiring a minimum 18-month nonverbal mental age for both the ADI-R and ADOS-2 (i.e., Modules 1–4) and independent walking for the ADOS-2 for valid administration.

ASD Classification. After reviewing all available data, the clinician assigned a consensus diagnosis of ASD or non-ASD and endorsed a degree of certainty on a scale of 1–5, where higher scores reflected greater clinical certainty. As noted above, in PHTS, administration of clinician-administered scales, i.e., the ADOS-2 and ADI-R, varied (Busch et al., 2019). As such, consensus diagnoses for the full PHTS sample reflect different data sources.

Analytic approach

Evaluation of the summary statistics within and across cohorts were used to accomplish the goals of this descriptive study. As comparison across groups was meant to inform the development of harmonized assessment strategies for phenotypically similar groups, rather than to test hypothesized differences across groups, statistical comparison was not performed. Visual inspection of the data determined whether mean and standard deviation (for normal distributions) or median and interquartile range (IQR) (for non-normal distributions) were used. The Supplementary Materials (https://doi.org/10.31234/osf.io/gqvr8_v1) detail the administration rates for each measure based on the standardization criteria outlined in the Measures section, but the summary statistics reflect only the valid administrations. Missingness was both informative (e.g., the participant did not meet the eligibility criteria for a given test, or could not achieve the minimum score possible on the test) and non-informative (e.g., related to logistical or administration errors). However, it was not possible to distinguish these cases in the database and so all missing data were treated equally.

Results

Demographic and Clinical Characteristics

INSERT TABLE 1

More than half of each group was male, and all groups were predominately White and non-Hispanic (Table 1). The mean (±SD) age of participants at enrollment across conditions was approximately 8 years (PMS: 8.41±4.59; TSC: 8.37±4.59, PHTS: 8.88±4.82). Seizures and sleep disturbances were among the most commonly reported medical conditions, and ASD and ADHD were the most commonly reported psychiatric diagnoses across all three conditions (Table 1).

Intellectual Functioning, Language, and Adaptive Functioning

INSERT TABLE 2

INSERT FIGURE 1

Intellectual Functioning. For all three conditions, the cognitive profile was relatively even across VIQ and NVIQ (Table 2). Most participants in the PMS sample were in the severe/profound range of ID, the majority of the TSC sample fell into the mild-to-moderate ID range, and the majority of the PHTS sample was in the borderline-to-average range of intellectual functioning (Figure 1C). However, visual inspection of the score distribution yielded important information about the impact of the combination of scores from multiple tests. There was a clear effect of test floor, resulting in distributional peaks at the floors of the Stanford Binet and DAS-II, while the MSEL DQ scores have a relatively normal distribution (Figure 1A).

Adaptive Functioning. Vineland-II caregiver interview (PMS) and caregiver survey (TSC/PHTS) standard scores indicated a degree of impairment similar to the cognitive scores, though the lower floor for the MSEL DQs versus standard scores allowed for a larger cognitive –

adaptive behavior discrepancy in the PMS sample (Table 2, Figure 1B). Within condition, the profiles were even, with similar scores across domains. However, all conditions exhibited a great deal of variability, with scores spanning the full range of impairment.

Language. Only one-third of the PMS participants were able to complete the PPVT-4 and EVT-2, compared to about two-thirds of the TSC sample and the majority of the PHTS sample (Table 2). Within all three conditions, the scores among those able to take the test were variable, ranging from extremely low to above average. For each condition, typical performance was similar for PPVT-4 and EVT-2, though the subsamples with valid data differed slightly.

Challenging Behavior

INSERT TABLE 3

Across conditions, clinical elevations in ABC scores were observed for over 40% of each cohort on the Stereotypy and Lethargy subdomains (Table 3). Clinical elevations in Irritability, a common treatment target in syndromic IDDs, was lowest in the PHTS cohort (32%) and highest in TSC (58%). Another common treatment target, Hyperactivity, was reported at higher rates in PMS (63%), followed by TSC (53%) and PHTS (30%).

The validity of the CBCL was limited by the rate of ID: fewer than 10 participants in the PMS sample, less than 1/3 of the TSC sample, and only about half of the PHTS received valid CBCL administrations. Limiting evaluation to the TSC and PHTS subsamples without ID, a high rate of clinically elevated Internalizing Symptoms scores was observed for both cohorts (TSC, 52%; PHTS, 41%). Externalizing Symptoms scores were clinically elevated for 41% of the TSC sample but only 12% of the PHTS sample.

ASD Symptoms and Diagnosis

INSERT TABLE 4

Scores on the RBS-R reflected low rates of repetitive behaviors and restricted interests in PMS, TSC, and PHTS evaluated for ASD. The SRS short-form scores are shown in Table 4, alongside T-scores for participants with NVIQ ≥ 70.

INSERT TABLE 5

Table 5 summarizes the results of ASD diagnostic evaluation. In the PMS and TSC studies, all participants were systematically assessed for ASD. A consensus diagnosis of ASD was common in both cohorts (PMS, 59%; TSC, 44%); the rate of parent-reported historical diagnosis was higher for PMS (69%) but similar for TSC (47%). The validity of the ADI-R and ADOS-2 were limited by the rate of profound ID in the PMS and TSC cohorts. Among those with valid assessment, the tools had good sensitivity (ADI-R: PMS, 88%; TSC, 83%; ADOS-2: PMS, 96%, TSC, 94%). While the specificity of both tools was good in the TSC cohort (ADI-R, 88%; ADOS-2, 91%), it was poor for PMS (ADI-R, 41%; ADOS-2, 65%).

The rate of consensus ASD diagnosis in PHTS was 62%. PHTS participants received comprehensive ASD evaluations based on clinical suspicion or history of diagnosis (87% of those assessed had a historical diagnosis); among these individuals, the sensitivity of the ADI-R and ADOS-2 was excellent (90%). However, the selective autism screening procedures for PHTS described in the Methods section limit interpretation of these psychometric data.

Discussion

The DSC is a network of large, multisite, natural history studies of the rare genetic conditions TSC, PMS, and PHTS. A primary goal of the DSC was clinical trial readiness, enabling nimble response to ongoing technological developments that may yield life-changing interventions for these conditions. A major aspect of clinical trial readiness is having identified the concepts of interest, or the aspects of the phenotype which affect how a person feels,

functions, or survives, as well as having a good understanding of how best to assess those concepts. However, this is particularly challenging for natural history studies of rare genetic conditions, where the populations are small and heterogeneous, and the developmental nature of relevant concepts means that their manifestation necessarily changes across development. In the current manuscript, we addressed these needs by identifying areas of clinical interest and describing the scope and limitations of conventional measurement strategies for these concepts in neurologically and medically complex genetic syndromes. The results of this project will inform trial planning for these conditions, as well as other rare genetic conditions affecting neurodevelopment that share similar phenotypic features.

Adaptive Functioning

Consistent with extant literature, ID was most common for participants in the PMS study and least common in the PHTS study. However, the full range of impairment was observed within all three conditions. Deficits in adaptive behavior, or the ability to perform the skills necessary to achieve an age-appropriate level of independence, are a core component of the ID diagnosis. Adaptive behavior is of clear clinical relevance, as it reflects the cumulative and global effects on functioning of the disease. As is common for genetic conditions affecting neurodevelopment, developmental delay, motor impairment, and communication impairments are among the top concerns for all three conditions in the DSC (e.g., Frazier et al., 2023; Gizzo et al., 2024; Ho et al., 2017; Landlust et al., 2023) – concepts which are all assessed by measures of adaptive behavior. In the DSC, adaptive behavior was assessed with the Vineland-II. Because there are no IQ or mental age restrictions on use of the Vineland-II, it was feasible for all participants in all three cohorts. Thus, adaptive behavior emerged as a key clinical concept and

candidate for future use in clinical trials. However, we note several issues for further consideration.

The DSC employed both the caregiver survey and the semi-structured interview form of the Vineland-II. While differences in administration (respondent and basal/ceiling rules) preclude their interchangeable use, there are several factors to consider in selecting the most appropriate form for a new study. First, clinician time may be reduced when using the survey form, though the manual (Sparrow et al., 2005) notes that a clinician should review the responses with the caregiver (p. 44). For populations where more severe disability is common, however, the manual cautions that "… parents often report enjoying the semi-structured interview and find it comforting to be able to describe what their child does [as in the semi-structured interview] rather than what he or she doesn't do [as in the survey]" (p. 11; parentheticals added).

Second, Vineland-II adaptive behavior scores may yield higher-than-expected standard scores compared to cognitive scores, especially for younger participants (Furnier et al., 2024). The role of age in the interpretation of the adaptive behavior concept is important to consider. Younger individuals are expected to do fewer things independently, and impairment is likely to become more apparent as an individual ages. Indeed, Vineland-II scores tend to decline over age in samples with ID (e.g., Sullivan et al. , 2022) – a phenomenon observed in the PMS sample (see Srivastava et al., in press). This has been partially addressed by additional items at lower levels of ability in the Vineland-3, leading to systematically lower scores on that instrument (Farmer et al., 2020), but the issue is fundamental to the concept and therefore remains. Additionally, the use of DQs for some participants with severe cognitive impairment exacerbated the issue since DQs do not have a floor (see PMS scores in Table 2).

Cognitive ability, the other concept core to the ID diagnosis, is also strongly aligned with parent/caregiver concerns and is de facto the concept used to estimate ID in clinical research. The heterogeneity in cognitive ability observed in the DSC creates considerable challenges for clinical trial design. The first major issue is the selection of a single clinical outcome assessment. Following a hierarchical testing method for severe-to-profound IDD populations described by Soorya et al. (2018), we planned to use traditional IQ testing where possible, substituting developmental testing where necessary. The advantage of this approach is that it allows for an estimate of cognitive functioning for the entirety of the sample, whereas using only traditional IQ tests would have resulted in a high rate of missing data – for the PMS sample, 40% required out-of-age-range testing and would have had missing data if only the Stanford Binet were used. While the rates were much lower for TSC (6%) and PHTS (2%), we note that any systematic missingness is unacceptable for clinical endpoints.

However, the results of this study highlighted a few limitations of this hierarchical approach. In order to combine estimates across tests, norm-referenced (or approximations) scores were used. But even when participants could achieve basal on a traditional IQ test, we observed significant floor effects for all three conditions (see Figure 1A). While floor effects may be tolerable in a diagnostic context where the goal is to identify impairment, floor effects in a clinical trial context obscure variability and reduce responsiveness to change (Farmer et al., 2022). For the participants who could not receive standard scores, we used DQs, which are considered analogous. However, the use of DQ is compromised by significant concerns about validity, especially for older and minimally verbal individuals (Ostrolenk & Courchesne, 2023). Other approaches, including Z-scores, could also be considered if norm referenced scores are required (Farmer, Thurm, et al., in press).

Communication is another clinically meaningful concept (Ho et al., 2017). The Vineland measures caregiver-reported functioning in the areas of receptive, expressive, and written communication, and as described above was the only measure applicable to the full range of age and ability in this study. However, we also explored the feasibility of direct assessment using receptive and expressive vocabulary tests (i.e., EVT-2, PPVT-4) and found these tests were appropriate only for samples with less severe ID. Thus, potential cross-syndrome treatment or phenotyping efforts focused on direct language assessment may require ability-based inclusion criteria and interpretation must be tempered by consideration of systematic missingness related to severe phenotypes in a particular domain (i.e., ID).

ASD Symptoms and Diagnosis

Consensus ASD diagnosis was common in all three cohorts (PMS, 59%; TSC, 44%; PHTS, 62%), marking it as a potential concept of interest. For PMS and TSC, where recruitment was unrelated to ASD status, a significant methodological strength was the systematic approach to the ASD diagnostic process. The diagnosis of ASD can be challenging in the context of ID, because many of the deficits required of an ASD diagnosis occur to some extent in all individuals with ID (Thurm et al., 2019). Here, we ensured that our direct assessment (ADOS-2) and observer/clinician rated assessment (ADI-R) were analyzed only in individuals meeting developmental and physical/sensory requirements of the measures. Importantly, this meant that these gold-standard assessments were not appropriate for half of the PMS sample and up to 30% of the TSC sample, limiting their utility in future clinical trials. Conditioned upon a nonverbal mental age of at least 18 months, the sensitivity of the ADOS-2 and ADI-R was good in both cohorts, but their specificity was poor for PMS – likely due to ID severity.

ASD diagnosis is less likely than a dimensional assessment of symptom severity to be part of a clinical trial endpoint. The DSC studies contained two such measures, the SRS-2 and the RBS-R. Both assessments are commonly used observer-reported assessments of ASD symptoms, but an important difference is their applicability for individuals with ID. The RBS-R was developed generally for IDD (versus ASD specifically) and the validity of the instrument was therefore unaffected by the rate of ID in the samples. Scores on the RBS-R were relatively low across conditions relative to standardization groups (Lam & Aman, 2007), indicating that the severity of behavior is not abnormal compared to general IDD and that RRB is not a unique clinical feature of the conditions, though it may still be meaningful. However, if repetitive behaviors present as clinically important concerns to families/caregivers, the low scores observed here indicate the RBS-R may not exhibit sufficient responsiveness in a clinical trial context.

The SRS, on the other hand, was not developed for use in individuals with ID. We addressed the variety of empirical evidence pointing to the lack of validity of the SRS when employed for individuals with ID, problem behavior, and/or limited language (Gergoudis et al., 2020; Havdahl et al., 2016) by using the 16-item short form (Sturm et al., 2017). The advantage of this approach is confidence that scores are more reflective of social communication behaviors than the participants' age, language level, cognitive ability, or challenging behaviors. A limitation of the short form is that no normative data exist, and so it is difficult to contextualize the scores. However, the mean scores in each condition indicates sufficient variability to detect a potential treatment effect, suggesting that the SRS short form may be a good candidate outcome measurement of social communication symptoms in predominately ID samples.

Challenging behaviors

Challenging behaviors can have a significant impact on quality of life and are rated by parents/caregivers as being an important area of concern (Ho et al., 2017), but differences from typically developing individuals in the function or cause of topographically similar behaviors complicates their assessment. Because individuals with ID were not included in the development of commonly used psychiatric screeners like the CBCL, domain-level scores are difficult to interpret, and normative data are likely not relevant. In the current study, we addressed this by using CBCL data only from participants with NVIQ scores ≥70. As a result, however, the CBCL was not considered valid for most PMS and TSC participants and half of PHTS participants, making it a poor choice for outcome measurement in future clinical trials.

Recognizing the limitations of the CBCL and knowing that measures validated in samples of individuals with IDD exist like the Developmental Behaviour Checklist (Einfeld & Tonge, 2002) and the ABC, we also employed the ABC. This scale is widely used in clinical trials, especially the Irritability subscale. Across all three conditions, mean scores were elevated in comparison to age- and sex-based normative data from special education settings but similar to scores from conditions with similar levels of ID severity (see Supplementary Table S5 in Miller et al., under review). For some subscales, such as Stereotypy or Inappropriate Speech, the numerically small mean scores suggest that responsivity might be limited in a clinical trial context, so if these are concepts of interest, the ABC may not be a good choice for outcome assessment.

Limitations and Future Directions

While there were considerable obstacles encountered in the outlined neurobehavioral approach, the careful, systematic cognitive and behavioral phenotyping protocol was crucial to systematically identifying limitations and future directions in selection of cognitive and

behavioral measurements for future studies. The detailed descriptive analysis here points to the importance of adapting protocols to account for heterogeneity within syndromic IDDs as well as within and across clinically meaningful concepts: cognition, communication, and behavioral challenges. Additionally, while psychometric data are a starting point for measurement selection in phenotyping and outcome protocols, they may require additional scrutiny when applied to special populations (Farmer et al., 2024; Gell et al., 2024). We also note an important limitation in the representativeness of the cohorts in DSC-1, which were predominantly White and non-Hispanic. This limitation is likely due to challenges with equity and accessibility in genetic testing and research recruitment generally (Cole et al., 2025), as well as study processes which were limited to English-speaking participants. Further, we are unable to analyze how specific medical comorbidities (e.g., epilepsy), significant life changes or onset of new psychiatric problems or significant regression of skills affect developmental trajectories since the sample included a wide age range of youth, and this analysis covers only the first time point.

We identified several commonly used measures that may be appropriate for use in future clinical trials across disease and ability levels (e.g., Vineland), but found significant limitations in commonly used instruments (e.g., CBCL), which may render them invalid for trials including samples with ID. While it would be helpful to the field for recommendations to be made from this or other natural history studies about which measures should be definitively considered or not from clinical trial use in varying severity levels of ID, there are many other considerations that need to be evaluated that may be trial or condition specific, including ages, timeframe, requirement for sensitivity to change and other features of the specific condition.

It is important to note that the current descriptive analysis is based exclusively on cross-sectional data, so we did not address sensitivity to change. For example, instead of the standard

scores described here, theory (Eisengart et al., 2022; Farmer et al., 2022) and quantitative data (Farmer et al., 2023, 2024; Kwok et al., 2022) support the use of person ability scores for clinical trial endpoints (e.g., Stanford Binet Change Sensitive Scores, DAS Ability Scores, or growth scales values for the EVT, PPVT and Vineland) for developmental domains like cognitive and communication. This is because ability scores are not subject to normative floor effects, are measured at the interval level, and are more responsive to change than normative scores. Crucially, however, the use of person ability scores is not compatible with a hierarchical approach to testing, because they cannot be compared across tests. These considerations led to our recommendation to prioritize adaptive functioning as a clinical trial endpoint, capitalizing on it as a well-established clinical outcome assessment that contain ability scores (i.e., Vineland-3) and minimal floor effects across age and ability levels.

While we identified several areas where clinically significant symptoms were observed, we note that our assessment of clinical significance was based on relative profiles from standardized datasets (e.g., in comparison to the general population or individuals with IDD). Patient and caregiver input on the meaningfulness of the measured concepts, the assessments, and input on what constitutes meaningful change in those assessments is an important and ongoing research area for genetic conditions associated with neurodevelopmental disorders (Connor-Ahmad et al., 2023; Downs et al., 2024; Gizzo et al., 2024; Hecker et al., 2024). The development of patient-reported outcome measures (Müller et al., 2023) may be used in tandem with standardized measures. Many of the symptoms observed to occur at high rates in these samples are known to be associated with poorer ratings of family quality of life, and so measures of the effects of these symptoms on the family quality of life require exploration.

In summary, the current report highlights foundational details on the application of commonly used cognitive/behavioral tools in studies of PMS, TSC, and PTEN and stresses important measurement, clinical, consumer, and contextual considerations to evaluate for clinical trial readiness.

References

Achenbach, T. M. (1999). The Child Behavior Checklist and related instruments. In *The use of psychological testing for treatment planning and outcomes assessment, 2nd ed.* (pp. 429–466). Lawrence Erlbaum Associates Publishers.

Aman, M. G., & Singh, N. N. (1994). *The Aberrant Behavior Checklist - Community Version*. Slosson Education Publications, Inc.

Aman, M. G., & Singh, N. N. (2017). *Aberrant behavior checklist manual (second edition)* (second). Slosson Educational Publications.

American Medical Association. (2018). *ICD-10-PCS 2019: the complete official codebook*.

Anderlid, B. M., Schoumans, J., Annerén, G., Tapia-Paez, I., Dumanski, J., Blennow, E., & Nordenskjöld, M. (2002). FISH-mapping of a 100-kb terminal 22q13 deletion. *Human Genetics*, *110*(5). https://doi.org/10.1007/s00439-002-0713-7

Betancur, C., & Buxbaum, J. D. (2013). SHANK3 haploinsufficiency: A “common” but underdiagnosed highly penetrant monogenic cause of autism spectrum disorders. *Molecular Autism*, *4*(1). https://doi.org/10.1186/2040-2392-4-17

Bishop, S. L., Guthrie, W., Coffing, M., & Lord, C. (2011). Convergent validity of the Mullen Scales of Early Learning and the Differential Ability Scales in children with autism spectrum disorders. *American Journal on Intellectual and Developmental Disabilities*, *116*(5). https://doi.org/10.1352/1944-7558-116.5.331

Bodfish, J. W., Symons, F. J., Parker, D. E., & Lewis, M. H. (2000). Varieties of repetitive behavior in autism: Comparisons to mental retardation. *Journal of Autism and Developmental Disorders*, *30*(3). https://doi.org/10.1023/A:1005596502855

Busch, R. M., Chapin, J. S., Mester, J., Ferguson, L., Haut, J. S., Frazier, T. W., & Eng, C. (2013). Cognitive characteristics of PTEN hamartoma tumor syndromes. *Genetics in Medicine*, *15*(7). https://doi.org/10.1038/gim.2013.1

Busch, R. M., Frazier, T. W., Sonneborn, C., Hogue, O., Klaas, P., Srivastava, S., Hardan, A. Y., Martinez-Agosto, J. A., Sahin, M., & Eng, C. (2023). Longitudinal neurobehavioral profiles in children and young adults with PTEN hamartoma tumor syndrome and reliable methods for assessing neurobehavioral change. *Journal of Neurodevelopmental Disorders*, *15*(1). https://doi.org/10.1186/s11689-022-09468-4

Busch, R. M., Srivastava, S., Hogue, O., Frazier, T. W., Klaas, P., Hardan, A., Martinez-Agosto, J. A., Sahin, M., Eng, C., & Consortium, D. S. (2019). Neurobehavioral phenotype of autism spectrum disorder associated with germline heterozygous mutations in PTEN. *Translational Psychiatry*, *9*(1), 251–253. https://doi.org/10.1038/s41398-019-0588-1 [doi]

Butler, M. G., Dazouki, M. J., Zhou, X. P., Talebizadeh, Z., Brown, M., Takahashi, T. N., Miles, J. H., Wang, C. H., Stratton, R., Pilarski, R., & Eng, C. (2005). Subset of individuals with autism spectrum disorders and extreme macrocephaly associated with germline PTEN tumour suppressor gene mutations. *Journal of Medical Genetics*, *42*(4). https://doi.org/10.1136/jmg.2004.024646

Ciaccio, C., Saletti, V., D'Arrigo, S., Esposito, S., Alfei, E., Moroni, I., Tonduti, D., Chiapparini, L., Pantaleoni, C., & Milani, D. (2019). Clinical spectrum of PTEN mutation in pediatric patients. A bicenter experience. *European Journal of Medical Genetics*, *62*(12). https://doi.org/10.1016/j.ejmg.2018.12.001

Cole, J. J., Williams, J. P., Sellitto, A. D., Baratta, L. R., Huecker, J. B., Baldridge, D., Kannampallil, T., Gurnett, C. A., & Balls-Berry, J. E. (2025). Association of Social Determinants of Health With Genetic Test Request and Completion Rates in Children With Neurologic Disorders. *Neurology*, *104*(5), e210275. https://doi.org/10.1212/WNL.0000000000210275

Connor-Ahmad, S., Tjeertes, J., Chladek, M., Newton, L., Symonds, T., Clinch, S., Vincenzi, B., & McDougall, F. (2023). Developing Angelman syndrome-specific clinician-reported and caregiver-reported measures to support holistic, patient-centered drug development. *Orphanet Journal of Rare Diseases*, *18*(1). https://doi.org/10.1186/s13023-023-02729-y

Constantino, J., & Gruber, C. (2005). *The Social Responsiveness Scale*. Western Psychological Services.

Crino, P. B., Nathanson, K. L., & Henske, E. P. (2006). Medical progress: The tuberous sclerosis complex. *The New England Journal of Medicine*, *355*(13). https://doi.org/10.1056/NEJMra055323

Cummings, K., Watkins, A., Jones, C., Dias, R., & Welham, A. (2022). Behavioural and psychological features of PTEN mutations: a systematic review of the literature and meta-analysis of the prevalence of autism spectrum disorder characteristics. In *Journal*

of Neurodevelopmental Disorders (Vol. 14, Issue 1). https://doi.org/10.1186/s11689-021-09406-w

Curatolo, P., Bombardieri, R., & Jozwiak, S. (2008). Tuberous sclerosis. *The Lancet*, *372*(9639), 657–668. https://doi.org/10.1016/S0140-6736(08)61279-9

Darnell, J. C., Van Driesche, S. J., Zhang, C., Hung, K. Y. S., Mele, A., Fraser, C. E., Stone, E. F., Chen, C., Fak, J. J., Chi, S. W., Licatalosi, D. D., Richter, J. D., & Darnell, R. B. (2011). FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell*, *146*(2). https://doi.org/10.1016/j.cell.2011.06.013

de Saint Martin, A., Napuri, S., & Nguyen, S. (2022). Tuberous sclerosis complex and epilepsy in infancy: prevention and early diagnosis. In *Archives de Pediatrie* (Vol. 29, Issue 5). https://doi.org/10.1016/S0929-693X(22)00284-6

de Vries, P. J., Heunis, T. M., Vanclooster, S., Chambers, N., Bissell, S., Byars, A. W., Flinn, J., Gipson, T. T., van Eeghen, A. M., Waltereit, R., Capal, J. K., Cukier, S., Davis, P. E., Smith, C., Kingswood, J. C., Schoeters, E., Srivastava, S., Takei, M., Gardner-Lubbe, S., … Jansen, A. C. (2023). International consensus recommendations for the identification and treatment of tuberous sclerosis complex-associated neuropsychiatric disorders (TAND). *Journal of Neurodevelopmental Disorders*, *15*(1). https://doi.org/10.1186/s11689-023-09500-1

Dekker, M. C., Koot, H. M., Van Der Ende, J., & Verhulst, F. C. (2002). Emotional and behavioral problems in children and adolescents with and without intellectual disability. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *43*(8). https://doi.org/10.1111/1469-7610.00235

Dhawan, A., Baitamouni, S., Liu, D., Yehia, L., Anthony, K., McCarther, A., Tischkowitz, M., MacFarland, S. P., Ngeow, J., Hoogerbrugge, N., & Eng, C. (2025). Cancer and Overgrowth Manifestations of PTEN Hamartoma Tumour Syndrome: Management Recommendations from the International PHTS Consensus Guidelines Working Group. *Clinical Cancer Research*. https://doi.org/10.1158/1078-0432.CCR-24-3819

Dille, Y., Lagae, L., Swillen, A., & Buggenhout, G. Van. (2022). Neurodevelopmental profile and stages of regression in Phelan–McDermid syndrome. *Developmental Medicine and Child Neurology*. https://doi.org/10.1111/dmcn.15482

Dölen, G., & Sahin, M. (2016). Editorial: Essential pathways and circuits of autism pathogenesis. In *Frontiers in Neuroscience* (Vol. 10, Issue APR). https://doi.org/10.3389/fnins.2016.00182

Dovgan, K., Mazurek, M. O., & Hansen, J. (2019). Measurement invariance of the child behavior checklist in children with autism spectrum disorder with and without intellectual disability: Follow-up study. *Research in Autism Spectrum Disorders*, *58*. https://doi.org/10.1016/j.rasd.2018.11.009

Downs, J., Ludwig, N. N., Wojnaroski, M., Keeley, J., Schust Myers, L., Chapman, C. A. T., Hecker, J. E., Conecker, G., & Berg, A. T. (2024). What does better look like in individuals with severe neurodevelopmental impairments? A qualitative descriptive study on SCN2A-related developmental and epileptic encephalopathy. *Quality of Life Research*, *33*(2). https://doi.org/10.1007/s11136-023-03543-6

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test--Fourth Edition (PPVT-4)*. Pearson.

Einfeld, S., & Tonge, B. (2002). *Manual for the Developmental Behaviour Checklist: Primary Carer Version (DBC-P) and Teacher Version (DBC-T) 2nd ed* (2nd ed.). Monash University Centre for Developmental Psychiatry and Psychology.

Eisengart, J. B., Daniel, M. H., Adams, H. R., Williams, P., Kuca, B., & Shapiro, E. (2022). Increasing precision in the measurement of change in pediatric neurodegenerative disease. In *Molecular Genetics and Metabolism* (Vol. 137, Issues 1–2). https://doi.org/10.1016/j.ymgme.2022.09.001

Elliott, C. D. (2007). *Differential ability scales* (2nd Edition). The psychological corporation.

Esbensen, A. J., Hoffman, E. K., Shaffer, R., Chen, E., Patel, L., & Jacola, L. (2018). Reliability of parent report measures of behaviour in children with Down syndrome. *Journal of Intellectual Disability Research*, *62*(9). https://doi.org/10.1111/jir.12533

Farach, L. S., Pearson, D. A., Woodhouse, J. P., Schraw, J. M., Sahin, M., Krueger, D. A., Wu, J. Y., Bebin, E. M., Lupo, P. J., Au, K. S., Northrup, H., Sahin, M., Krueger, D., Bebin, M., Wu, J. Y., Northrup, H., Warfield, S., Peters, J., Scherrer, B., … Murray, D. S. (2019). Tuberous Sclerosis Complex Genotypes and Developmental Phenotype. *Pediatric Neurology*, *96*. https://doi.org/10.1016/j.pediatrneurol.2019.03.003

Farmer, C., Adedipe, D., Bal, V. H., Chlebowski, C., & Thurm, A. (2020). Concordance of the Vineland Adaptive Behavior Scales, second and third editions. *Journal of Intellectual Disability Research*, *64*(1). https://doi.org/10.1111/jir.12691

Farmer, C., Giserman-Kiss, I., Mohanty, E., Soorya, L., Sahin, M., Kolevzon, A., Buxbaum, J., Berry-Kravis, E., Powell, C., Bernstein, J., & Thurm, A. (n.d.). Retrospective

reports of skill attainment and loss in Phelan-McDermid Syndrome (under review). *American Journal on Intellectual and Developmental Disabilities*.

Farmer, C., Golden, C., & Thurm, A. (2016). Concurrent validity of the differential ability scales, second edition with the Mullen Scales of Early Learning in young children with and without neurodevelopmental disorders. *Child Neuropsychology*, *22*(5). https://doi.org/10.1080/09297049.2015.1020775

Farmer, C., Kaat, A. J., Berry-Kravis, E., & Thurm, A. (2022). Psychometric perspectives on developmental outcome and endpoint selection in treatment trials for genetic conditions associated with neurodevelopmental disorder. In *International Review of Research in Developmental Disabilities* (Vol. 62). https://doi.org/10.1016/bs.irrdd.2022.05.001

Farmer, C., Thurm, A., Das, T., Bebin, E. M., Bernstein, J., Berry-Kravis, E., Buxbaum, J. D., & et al. (2024). Which score for what? Operationalizing standardized cognitive test performance for the assessment of change. *American Journal on Intellectual and Developmental Disabilities*.

Farmer, C., Thurm, A., Troy, J. D., & Kaat, A. J. (2023). Comparing ability and norm-referenced scores as clinical trial outcomes for neurodevelopmental disabilities: a simulation study. *Journal of Neurodevelopmental Disorders*, *15*(1). https://doi.org/10.1186/S11689-022-09474-6

Food and Drug Administration. (2019, March). *Rare Diseases: Natural History Studies for Drug Development*. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/rare-diseases-natural-history-studies-drug-development

Food and Drug Administration. (2022, June). *Patient-Focused Drug Development: Selecting, Developing, or Modifying Fit-for-Purpose Clinical Outcome Assessments*. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-focused-drug-development-selecting-developing-or-modifying-fit-purpose-clinical-outcome

Frank, Y. (2021). The Neurological Manifestations of Phelan-McDermid Syndrome. *Pediatric Neurology*, *122*, 59–64. https://doi.org/S0887-8994(21)00124-7 [pii]

Frazier, T. W., Busch, R. M., Klaas, P., Lachlan, K., Jeste, S., Kolevzon, A., Loth, E., Harris, J., Speer, L., Pepper, T., Anthony, K., Graglia, J. M., Delagrammatikas, C., Bedrosian-Sermone, S., Beekhuyzen, J., Smith-Hicks, C., Sahin, M., Eng, C., Hardan, A. Y., & Uljarević, M. (2023). Development of informant-report neurobehavioral survey scales for PTEN hamartoma tumor syndrome and related neurodevelopmental genetic syndromes. *American Journal of Medical Genetics, Part A*, *191*(7). https://doi.org/10.1002/ajmg.a.63195

Frazier, T. W., Embacher, R., Tilot, A. K., Koenig, K., Mester, J., & Eng, C. (2015). Molecular and phenotypic abnormalities in individuals with germline heterozygous PTEN mutations and autism. *Molecular Psychiatry*, *20*(9). https://doi.org/10.1038/mp.2014.125

Furnier, S. M., Gangnon, R., Daniels, J. L., Ellis Weismer, S., Nadler, C., Pazol, K., Reyes, N. M., Rosenberg, S., Rubenstein, E., Wiggins, L. D., Yeargin-Allsopp, M., & Durkin, M. S. (2024). Racial and ethnic disparities in the co-occurrence of intellectual

disability and autism: Impact of incorporating measures of adaptive functioning. *Autism Research*, *17*(3). https://doi.org/10.1002/aur.3107

Gell, M., Eickhoff, S. B., Omidvarnia, A., Küppers, V., Patil, K. R., Satterthwaite, T. D., Müller, V. I., & Langner, R. (2024). How measurement noise limits the accuracy of brain-behaviour predictions. *Nature Communications*, *15*(1), 10678. https://doi.org/10.1038/s41467-024-54022-6

Gergoudis, K., Weinberg, A., Templin, J., Farmer, C., Durkin, A., Weissman, J., Siper, P., Foss-Feig, J., del Pilar Trelles, M., Bernstein, J. A., Buxbaum, J. D., Berry-Kravis, E., Powell, C. M., Sahin, M., Soorya, L., Thurm, A., & Kolevzon, A. (2020). Psychometric Study of the Social Responsiveness Scale in Phelan–McDermid Syndrome. *Autism Research*, *13*(8). https://doi.org/10.1002/aur.2299

Gizzo, L., Bliss, G., Palaty, C., & Kolevzon, A. (2024). Caregiver perspectives on patient-focused drug development for Phelan-McDermid syndrome. *Orphanet Journal of Rare Diseases*, *19*(1). https://doi.org/10.1186/s13023-024-03141-w

Goh, S., Kwiatkowski, D. J., Dorer, D. J., & Thiele, E. A. (2005). Infantile spasms and intellectual outcomes in children with tuberous sclerosis complex. *Neurology*, *65*(2). https://doi.org/10.1212/01.wnl.0000168908.78118.99

Hansen-Kiss, E., Beinkampen, S., Adler, B., Frazier, T., Prior, T., Erdman, S., Eng, C., & Herman, G. (2017). A retrospective chart review of the features of PTEN hamartoma tumour syndrome in children. *Journal of Medical Genetics*, *54*(7). https://doi.org/10.1136/jmedgenet-2016-104484

Havdahl, K. A., Hus Bal, V., Huerta, M., Pickles, A., Øyen, A. S., Stoltenberg, C., Lord, C., & Bishop, S. L. (2016). Multidimensional Influences on Autism Symptom Measures: Implications for Use in Etiological Research. *Journal of the American Academy of Child and Adolescent Psychiatry*, *55*(12). https://doi.org/10.1016/j.jaac.2016.09.490

Hecker, J., Conecker, G., Chapman, C., Hommer, R., Ludwig, N. N., Sevinc, G., Te, S., Wojnaroski, M., Downs, J., & Berg, A. T. (2024). Patient-advocate-led global coalition adapting fit-for-purpose outcomes measures to assure meaningful inclusion of DEEs in clinical trials. https://doi.org/10.1177/26330040241249762. *Therapeutic Advances in Rare Disease*, *18*.

Ho, C. N., Rushing, G., Valentine, J. E., Rosbeck, K. L., & Roberds, S. L. (2017). *The Voice of the Patient: A Report from the Tuberous Sclerosis Alliance's Externally-Led Patient-Focused Drug Development Meeting*. https://www.tscalliance.org/wp-content/uploads/2018/01/Voice-of-the-Patient-Tuberous-Sclerosis-Alliance.pdf

Hobert, J. A., & Eng, C. (2009). PTEN hamartoma tumor syndrome: An overview. In *Genetics in Medicine* (Vol. 11, Issue 10). https://doi.org/10.1097/GIM.0b013e3181ac9aea

Hus, V., Bishop, S., Gotham, K., Huerta, M., & Lord, C. (2013). Factors influencing scores on the social responsiveness scale. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *54*(2). https://doi.org/10.1111/j.1469-7610.2012.02589.x

Jansen, F. E., Braams, O., Vincken, K. L., Algra, A., Anbeek, P., Jennekens-Schinkel, A., Halley, D., Zonnenberg, B. A., Van Den Ouweland, A., Van Huffelen, A. C., Van Nieuwenhuizen, O., & Nellist, M. (2008). Overlapping neurologic and cognitive

phenotypes in patients with TSC1 or TSC2 mutations. *Neurology, 70*(12). https://doi.org/10.1212/01.wnl.0000280578.99900.96

Joinson, C., O'Callaghan, F. J., Osborne, J. P., Martyn, C., Harris, T., & Bolton, P. F. (2003). Learning disability and epilepsy in an epidemiological sample of individuals with tuberous sclerosis complex. *Psychological Medicine, 33*(2). https://doi.org/10.1017/S0033291702007092

Kidd, S. A., Berry-Kravis, E., Choo, T. H., Chen, C., Esler, A., Hoffmann, A., Andrews, H. F., & Kaufmann, W. E. (2020). Improving the Diagnosis of Autism Spectrum Disorder in Fragile X Syndrome by Adapting the Social Communication Questionnaire and the Social Responsiveness Scale-2. *Journal of Autism and Developmental Disorders, 50*(9). https://doi.org/10.1007/s10803-019-04148-0

Kim, S. H., Thurm, A., Shumway, S., & Lord, C. (2013). Multisite study of new autism diagnostic interview-revised (ADI-R) Algorithms for toddlers and young preschoolers. *Journal of Autism and Developmental Disorders, 43*(7). https://doi.org/10.1007/s10803-012-1696-4

Kohlenberg, T. M., Trelles, M. P., McLarney, B., Betancur, C., Thurm, A., & Kolevzon, A. (2020). Psychiatric illness and regression in individuals with Phelan-McDermid syndrome. *Journal of Neurodevelopmental Disorders, 12*(1). https://doi.org/10.1186/s11689-020-9309-6

Kolevzon, A., Delaby, E., Berry-Kravis, E., Buxbaum, J. D., & Betancur, C. (2019). Neuropsychiatric decompensation in adolescents and adults with Phelan-McDermid

syndrome: A systematic review of the literature. In *Molecular Autism* (Vol. 10, Issue 1). https://doi.org/10.1186/s13229-019-0291-3

Koskentausta, T., Iivanainen, M., & Almqvist, F. (2004). CBCL in the assessment of psychopathology in Finnish children with intellectual disability. *Research in Developmental Disabilities*, *25*(4). https://doi.org/10.1016/j.ridd.2003.12.001

Krueger, D. A., Northrup, H., Krueger, D. A., Roberds, S., Smith, K., Sampson, J., Korf, B., Kwiatkowski, D. J., Mowat, D., Nellist, M., Povey, S., de Vries, P., Byars, A., Dunn, D., Ess, K., Hook, D., Jansen, A., King, B., Sahin, M., … Frost, M. D. (2013). Tuberous sclerosis complex surveillance and management: Recommendations of the 2012 international tuberous sclerosis complex consensus conference. *Pediatric Neurology*, *49*(4). https://doi.org/10.1016/j.pediatrneurol.2013.08.002

Kwok, E., Feiner, H., Grauzer, J., Kaat, A., & Roberts, M. Y. (2022). Measuring Change During Intervention Using Norm-Referenced, Standardized Measures: A Comparison of Raw Scores, Standard Scores, Age Equivalents, and Growth Scale Values from the Preschool Language Scales–Fifth Edition. *Journal of Speech, Language, and Hearing Research*, *65*(11). https://doi.org/10.1044/2022_JSLHR-22-00122

Lam, K. S. L., & Aman, M. G. (2007). The repetitive behavior scale-revised: Independent validation in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *37*(5). https://doi.org/10.1007/s10803-006-0213-z

Landlust, A. M., Koza, S. A., Carbin, M., Walinga, M., Robert, S., Cooke, J., Vyshka, K., van Balkom, I. D. C., & van Ravenswaaij-Arts, C. (2023). Parental perspectives on Phelan-McDermid syndrome: Results of a worldwide survey. *European Journal of*

Medical Genetics, *66*(7), 104771. https://doi.org/https://doi.org/10.1016/j.ejmg.2023.104771

Levine, A., Davis, P., Zhang, B., Peters, J., Filip-Dhima, R., Warfield, S. K., Prohl, A., Capal, J., Krueger, D., Bebin, E. M., Northrup, H., Wu, J. Y., & Sahin, M. (2023). Epilepsy Severity Is Associated with Head Circumference and Growth Rate in Infants with Tuberous Sclerosis Complex. *Pediatric Neurology*, *144*, 26–32. https://doi.org/10.1016/j.pediatrneurol.2023.03.015

Levy, T., Farmer, C., Srivastava, S., Johnson, K., Trayvick, J., Brune, C., Massa, A., Silver, H., Siper, P. M., Zweifach, J., Halpern, D., Foss-Feig, J. H., Bernstein, J. A., Berry-Kravis, E., Powell, C. M., Sahin, M., Soorya, L. V., Thurm, A., Buxbaum, J. D., & Kolevzon, A. (2024). Genetic subtypes of Phelan-McDermid syndrome exhibit similar rates of change despite differences in level of impairment in developmental constructs. *American Journal on Intellectual and Developmental Disabilities*.

Levy, T., Foss-Feig, J. H., Betancur, C., Siper, P. M., Trelles-Thorne, M. D. P., Halpern, D., Frank, Y., Lozano, R., Layton, C., Britvan, B., Bernstein, J. A., Buxbaum, J. D., Berry-Kravis, E., Powell, C. M., Srivastava, S., Sahin, M., Soorya, L., Thurm, A., & Kolevzon, A. (2022). Strong evidence for genotype-phenotype correlations in Phelan-McDermid syndrome: Results from the developmental synaptopathies consortium. *Human Molecular Genetics*, *31*(4). https://doi.org/10.1093/hmg/ddab280

Levy, T., Gluckman, J., Siper, P. M., Halpern, D., Zweifach, J., Filip-Dhima, R., Holder, J., Trelles, M. P., Johnson, K., Bernstein, J. A., Berry-Kravis, E., Powell, C. M., Soorya, L. V., Thurm, A., Buxbaum, J. D., Sahin, M., Kolevzon, A., & Srivastava, S. (2024).

Clinical, genetic, and cognitive correlates of seizure occurrences in Phelan-McDermid syndrome. *Journal of Neurodevelopmental Disorders*, *16*(1), 25–12.

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2)*. Western Psychological Services.

Lord, C., Storoschuk, S., Rutter, M., & Pickles, A. (1993). Using the ADI-R to diagnose autism in preschool children. *Infant Mental Health Journal*, *14*(3). https://doi.org/10.1002/1097-0355(199323)14:3<234::AID-IMHJ2280140308>3.0.CO;2-F

Lyall, K., Hosseini, M., Ladd-Acosta, C., Ning, X., Catellier, D., Constantino, J. N., Croen, L. A., Kaat, A. J., Botteron, K., Bush, N. R., Dager, S. R., Duarte, C. S., Fallin, M. D., Hazlett, H., Hertz-Picciotto, I., Joseph, R. M., Karagas, M. R., Korrick, S., Landa, R., … Outcomes, program collaborators for E. influences on C. H. (2021). Distributional Properties and Criterion Validity of a Shortened Version of the Social Responsiveness Scale: Results from the ECHO Program and Implications for Social Communication Research. *Journal of Autism and Developmental Disorders*, *51*(7), 2241–2253. https://doi.org/10.1007/s10803-020-04667-1 [doi]

Maehama, T., & Dixon, J. E. (1999). PTEN: A tumour suppressor that functions as a phospholipid phosphatase. In *Trends in Cell Biology* (Vol. 9, Issue 4). https://doi.org/10.1016/S0962-8924(99)01519-6

Maehama, T., Taylor, G. S., & Dixon, J. E. (2001). PTEN and myotubularin: Novel phosphoinositide phosphatases. In *Annual Review of Biochemistry* (Vol. 70). https://doi.org/10.1146/annurev.biochem.70.1.247

Marshburn, E. C., & Aman, M. G. (1992). Factor validity and norms for the aberrant behavior checklist in a community sample of children with mental retardation. *Journal of Autism and Developmental Disorders; J Autism Dev Disord*, *22*(3), 357–373. https://doi.org/10.1007/BF01048240

Miller, J., Farmer, C., Blair, S., Bianconi, S., Akshoomoff, N., Anslem, I., ..., & Thurm, A. (n.d.). *Longitudinal Characterization of Males with X-Linked Creatine Transporter Deficiency: Final Results of the Vigilan Observational Study*.

Mullen, E. M. (1995). *Mullen Scales of Early Learning* (AGS ed.). American Guidance Service Inc.

Müller, A. R., Luijten, M. A. J., Haverman, L., de Ranitz-Greven, W. L., Janssens, P., Rietman, A. B., ten Hoopen, L. W., de Graaff, L. C. G., de Wit, M. C., Jansen, A. C., Gipson, T., Capal, J. K., de Vries, P. J., & van Eeghen, A. M. (2023). Understanding the impact of tuberous sclerosis complex: development and validation of the TSC-PROM. *BMC Medicine*, *21*(1). https://doi.org/10.1186/s12916-023-03012-4

National Organization of Rare Disorders. (2023, November 7). *PTEN Hamartoma Tumor Syndrome*. https://rarediseases.org/rare-diseases/pten-hamartoma-tumor-syndrome/#symptoms

Nisar, S., Bhat, A. A., Masoodi, T., Hashem, S., Akhtar, S., Ali, T. A., Amjad, S., Chawla, S., Bagga, P., Frenneaux, M. P., Reddy, R., Fakhro, K., & Haris, M. (2022). Genetics

of glutamate and its receptors in autism spectrum disorder. In *Molecular Psychiatry* (Vol. 27, Issue 5). https://doi.org/10.1038/s41380-022-01506-w

Northrup, H., Aronow, M. E., Bebin, E. M., Bissler, J., Darling, T. N., de Vries, P. J., Frost, M. D., Fuchs, Z., Gosnell, E. S., Gupta, N., Jansen, A. C., Jóźwiak, S., Kingswood, J. C., Knilans, T. K., McCormack, F. X., Pounders, A., Roberds, S. L., Rodriguez-Buritica, D. F., Roth, J., … Young, L. (2021). Updated International Tuberous Sclerosis Complex Diagnostic Criteria and Surveillance and Management Recommendations. *Pediatric Neurology, 123*. https://doi.org/10.1016/j.pediatrneurol.2021.07.011

Oberman, L. M., Boccuto, L., Cascio, L., Sarasua, S., & Kaufmann, W. E. (2015). Autism spectrum disorder in Phelan-McDermid syndrome: Initial characterization and genotype-phenotype correlations. *Orphanet Journal of Rare Diseases, 10*(1). https://doi.org/10.1186/s13023-015-0323-9

O'Callaghan, F. J. K., Shiell, A. W., Osborne, J. P., & Martyn, C. N. (1998). Prevalence of tuberous sclerosis estimated by capture-recapture analysis. *Lancet, 351*(9114). https://doi.org/10.1016/S0140-6736(05)78872-3

Ostrolenk, A., & Courchesne, V. (2023). Examining the validity of the use of ratio IQs in psychological assessments. *Acta Psychologica, 240*. https://doi.org/10.1016/j.actpsy.2023.104054

Pandolfi, V., Magyar, C. I., & Dill, C. A. (2009). Confirmatory factor analysis of the Child Behavior Checklist 1.5-5 in a sample of children with autism spectrum disorders.

Journal of Autism and Developmental Disorders, *39*(7). https://doi.org/10.1007/s10803-009-0716-5

Phelan, K., & McDermid, H. E. (2012). The 22q13.3 deletion syndrome (Phelan-McDermid syndrome). *Molecular Syndromology*, *2*(3–5). https://doi.org/10.1159/000334260

Piergies, A. M. H., Hirota, T., Monden, R., & Zheng, S. (2022). Subgrouping school-aged children on the autism spectrum based on co-occurring psychiatric symptoms. *Research in Autism Spectrum Disorders*, *95*. https://doi.org/10.1016/j.rasd.2022.101983

Risi, S., Lord, C., Gotham, K., Corsello, C., Chrysler, C., Szatmari, P., Cook, E. H., Leventhal, B. L., & Pickles, A. (2006). Combining information from multiple sources in the diagnosis of autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*(9). https://doi.org/10.1097/01.chi.0000227880.42780.0e

Roid, G. H., & Pomplun, M. (2012). *The Stanford-Binet Intelligence Scales* (5th Edition). Riverside Publishing.

Rutter, M., Le Couteur, A., & Lord, C. (2003). *ADI-R Autism Diagnostic Interview Revised*. Western Psychological Services.

Sahin, M., & Sur, M. (2015). Genes, circuits, and precision therapies for autism and related neurodevelopmental disorders. In *Science* (Vol. 350, Issue 6263). https://doi.org/10.1126/science.aab3897

Sakai, Y., Shaw, C. A., Dawson, B. C., Dugas, D. V., Al-Mohtaseb, Z., Hill, D. E., & Zoghbi, H. Y. (2011). Protein interactome reveals converging molecular pathways among autism disorders. *Science Translational Medicine*, *3*(86). https://doi.org/10.1126/scitranslmed.3002166

Sala, C., Piëch, V., Wilson, N. R., Passafaro, M., Liu, G., & Sheng, M. (2001). Regulation of dendritic spine morphology and synaptic function by Shank and Homer. *Neuron*, *31*(1). https://doi.org/10.1016/S0896-6273(01)00339-7

Sanders, S. J., Sahin, M., Hostyk, J., Thurm, A., Jacquemont, S., Avillach, P., Douard, E., Martin, C. L., Modi, M. E., Moreno-De-Luca, A., Raznahan, A., Anticevic, A., Dolmetsch, R., Feng, G., Geschwind, D. H., Glahn, D. C., Goldstein, D. B., Ledbetter, D. H., Mulle, J. G., … Bearden, C. E. (2019). A framework for the investigation of rare genetic disorders in neuropsychiatry. *Nature Medicine*, *25*(10). https://doi.org/10.1038/s41591-019-0581-5

Sarasua, S. M., Dwivedi, A., Boccuto, L., Rollins, J. D., Chen, C. F., Rogers, R. C., Phelan, K., Dupont, B. R., & Collins, J. S. (2011). Association between deletion size and important phenotypes expands the genomic region of interest in Phelan-McDermid syndrome (22q13 deletion syndrome). *Journal of Medical Genetics*, *48*(11). https://doi.org/10.1136/jmedgenet-2011-100225

Scalisi, F. C., Callea, M., Martinelli, D., Willoughby, C. E., Tadich, A. C., Castillo, M. A., Lacruz-Rengel, M. A., Medina, M., Grimaldi, P., Bertini, E., & Nevado, J. (2022). Clinical and Genetic Aspects of Phelan–McDermid Syndrome: An Interdisciplinary

Approach to Management. In *Genes* (Vol. 13, Issue 3). https://doi.org/10.3390/genes13030504

Schön, M., Lapunzina, P., Nevado, J., Mattina, T., Gunnarsson, C., Hadzsiev, K., Verpelli, C., Bourgeron, T., Jesse, S., van Ravenswaaij-Arts, C. M. A., & Hennekam, R. C. (2023). Definition and clinical variability of SHANK3-related Phelan-McDermid syndrome. *European Journal of Medical Genetics*, *66*(7). https://doi.org/10.1016/j.ejmg.2023.104754

Soorya, L., Kolevzon, A., Zweifach, J., Lim, T., Dobry, Y., Schwartz, L., Frank, Y., Wang, A. T., Cai, G., Parkhomenko, E., Halpern, D., Grodberg, D., Angarita, B., Willner, J. P., Yang, A., Canitano, R., Chaplin, W., Betancur, C., & Buxbaum, J. D. (2013). Prospective investigation of autism and genotype-phenotype correlations in 22q13 deletion syndrome and SHANK3 deficiency. *Molecular Autism*, *4*(1). https://doi.org/10.1186/2040-2392-4-18

Soorya, L., Leon, J., Trelles, M. P., & Thurm, A. (2018). Framework for assessing individuals with rare genetic disorders associated with profound intellectual and multiple disabilities (PIMD): the example of Phelan McDermid Syndrome. In *Clinical Neuropsychologist* (Vol. 32, Issue 7). https://doi.org/10.1080/13854046.2017.1413211

Sparrow, S. S., Balla, D. A., Cicchetti, D. V, & Doll, E. A. (Edgar A. (2005). Vineland-II: Vineland adaptive behavior scales: survey forms manual. In *Vineland adaptive behavior scales : survey forms manual* (2nd ed.). AGS Publishing.

Specchio, N., Pietrafusa, N., Trivisano, M., Moavero, R., De Palma, L., Ferretti, A., Vigevano, F., & Curatolo, P. (2020). Autism and Epilepsy in Patients with Tuberous

Sclerosis Complex. *Frontiers in Neurology*, *11*. https://doi.org/10.3389/fneur.2020.00639

Srivastava, S., Johnson, K., Farmer, C., Levy, T., Thurm, A., Soorya, L., Filip-Dhima, R., Quinlan, A., Bernstein, J. A., Berry-Kravis, E., Powell, C. M., Buxbaum, J. D., Sahin, M., & Kolevzon, A. (n.d.). Longitudinal trajectory of adaptive skills in Phelan-McDermid Syndrome. *American Journal on Intellectual and Developmental Disabilities*.

Srivastava, S., Sahin, M., Buxbaum, J. D., Berry-Kravis, E., Soorya, L. V., Thurm, A., Bernstein, J. A., Asante-Otoo, A., Bennett, W. E., Betancur, C., Brickhouse, T. H., Passos Bueno, M. R., Chopra, M., Christensen, C. K., Cully, J. L., Dies, K., Friedman, K., Gummere, B., Holder, J. L., … Kolevzon, A. (2023). Updated consensus guidelines on the management of Phelan–McDermid syndrome. In *American Journal of Medical Genetics, Part A* (Vol. 191, Issue 8). https://doi.org/10.1002/ajmg.a.63312

Sturm, A., Kuhfeld, M., Kasari, C., & McCracken, J. T. (2017). Development and validation of an item response theory-based Social Responsiveness Scale short form. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *58*(9), 1053–1061. https://doi.org/10.1111/jcpp.12731 [doi]

Sullivan, J., Wirrell, E., Knupp, K. G., Chen, D., Zafar, M., Flamini, R., Stutely, J., Brathwaite, C., Ventola, P., Avendaño, J., Parkerson, K. A., Wyant, N., & Ticho, B. (2022). Interim results of adaptive functioning and neurodevelopment in BUTTERFLY – An observational study of children and adolescents with Dravet syndrome. *Epilepsy and Behavior*, *137*. https://doi.org/10.1016/j.yebeh.2022.108955

Thurm, A., Farmer, C., Salzman, E., Lord, C., & Bishop, S. (2019). State of the field: Differentiating intellectual disability from autism spectrum disorder. In *Frontiers in Psychiatry* (Vol. 10). https://doi.org/10.3389/fpsyt.2019.00526

Tye, C., Thomas, L. E., Sampson, J. R., Lewis, J., O'Callaghan, F., Yates, J. R. W., & Bolton, P. F. (2018). Secular changes in severity of intellectual disability in tuberous sclerosis complex: A reflection of improved identification and treatment of epileptic spasms? *Epilepsia Open*, *3*(2). https://doi.org/10.1002/epi4.12111

van Eeghen, A. M., Stemkens, D., Fernández-Fructuoso, J. R., Maruani, A., Hadzsiev, K., Gaasterland, C. M. W., Klein Haneveld, M. J., Vyshka, K., Hugon, A., van Eeghen, A. M., Alhambra, N., Anderlid, B.-M., Andres, S., Aten, E., Guedes, R. B., Bonaglia, M. C., Bourgeron, T., Burdeus-Olavarrieta, M., Carbin, M. J., … van Balkom, I. D. C. (2023). Consensus recommendations on organization of care for individuals with Phelan-McDermid syndrome. *European Journal of Medical Genetics*, *66*(7), 104747. https://doi.org/https://doi.org/10.1016/j.ejmg.2023.104747

van Ravenswaaij-Arts, C. M. A., van Balkom, I. D. C., Jesse, S., & Bonaglia, M. C. (2023). Editorial: Towards a European consensus guideline for Phelan-McDermid syndrome. In *European Journal of Medical Genetics* (Vol. 66, Issue 5). https://doi.org/10.1016/j.ejmg.2023.104736

Varga, E. A., Pastore, M., Prior, T., Herman, G. E., & McBride, K. L. (2009). The prevalence of PTEN mutations in a clinical pediatric cohort with autism spectrum disorders, developmental delay, and macrocephaly. *Genetics in Medicine*, *11*(2). https://doi.org/10.1097/GIM.0b013e31818fd762

Williams, K. T. (2007). *Expressive vocabulary test-second edition*. NCS Pearson.

Wilson, H. L., Wong, A. C. C., Shaw, S. R., Tse, W. Y., Stapleton, G. A., Phelan, M. C., Hu, S., Marshall, J., & McDermid, H. E. (2003). Molecular characterisation of the 22q13 deletion syndrome supports the role of haploinsufficiency of SHANK3/PROSAP2 in the major neurological symptoms. *Journal of Medical Genetics*, *40*(8). https://doi.org/10.1136/jmg.40.8.575

Winden, K. D., Ebrahimi-Fakhari, D., & Sahin, M. (2018). Abnormal mTOR Activation in Autism. In *Annual Review of Neuroscience* (Vol. 41). https://doi.org/10.1146/annurev-neuro-080317-061747

Xu, N., Lv, H., Yang, T., Du, X., Sun, Y., Xiao, B., Fan, Y., Luo, X., Zhan, Y., Wang, L., Li, F., & Yu, Y. (2020). A 29 Mainland Chinese cohort of patients with Phelan–McDermid syndrome: genotype–phenotype correlations and the role of SHANK3 haploinsufficiency in the important phenotypes. *Orphanet Journal of Rare Diseases*, *15*(1). https://doi.org/10.1186/s13023-020-01592-5

Yum, M. S., Lee, E. H., & Ko, T. S. (2013). Vigabatrin and mental retardation in tuberous sclerosis: Infantile spasms versus focal seizures. *Journal of Child Neurology*, *28*(3). https://doi.org/10.1177/0883073812446485

Figure Legend

Figure 1. Intellectual Functioning and Adaptive Functioning in DSC-1 cohort. (A) Nonverbal IQ/DQ (developmental quotient) and (B) Vineland-2 Adaptive Behavior Composite (ABC) standard score distributions. (C) Nonverbal IQ/DQ (NVIQ) categorized according to ICD 10.

Figure 1

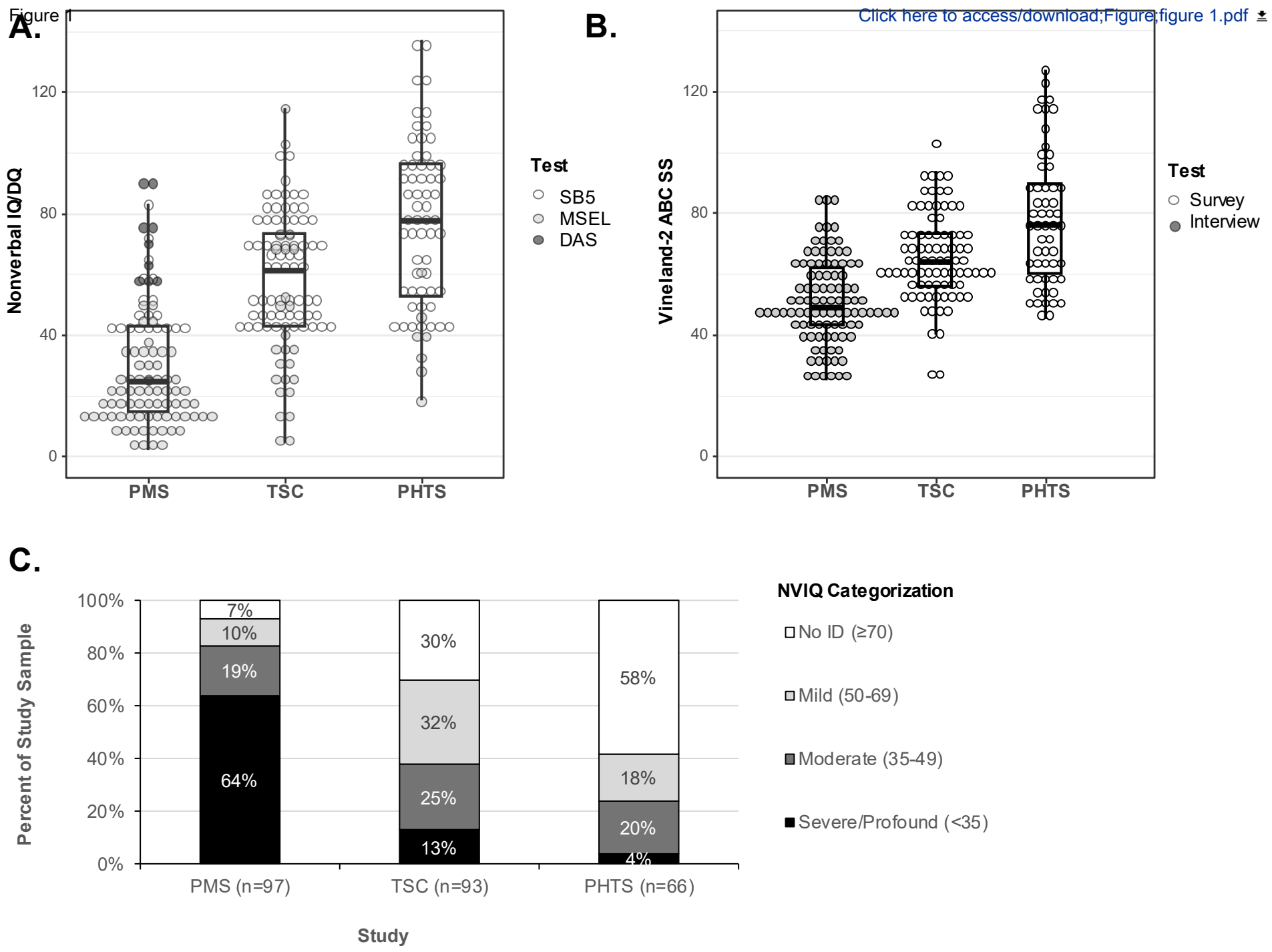

Table 1

Participant Demographic and Clinical Characteristics

Characteristic	PMS (N=98)	TSC (N=98)	PHTS (N=69)
	n (%)	n (%)	n (%)
Gender			
Female	45 (46)	37 (38)	19 (27)
Male	53 (54)	61 (62)	50 (73)
Race			
American Indian or Alaskan Native	1 (1)	0	0
Asian	8 (8)	1 (1)	2 (3)
Black or African American	2 (2)	3 (3)	1 (2)
White	83 (85)	84 (86)	54 (78)
More than one race	1 (1)	8 (8)	9 (13)
Unknown/not reported	3 (3)	2 (2)	3 (4)
Ethnicity			
Hispanic	11 (11)	16 (16)	7 (10)
Non-Hispanic	84 (86)	80 (82)	60 (87)
Unknown/not reported	3 (3)	2 (2)	2 (3)
Caregiver reported medical history	98	92	64
History of seizures	35 (36)	87 (95)	11 (17)
Disrupted sleep	46 (48)	37 (40)	23 (36)
Deafness	3 (3)	1 (1)	4 (6)
Significant vision loss	2 (2)	5 (6)	1 (2)
Significant motor impairment (not walking)	2 (2)	0	0
Caregiver reported neurodevelopmental psychiatric history	98	92	64
Autism Spectrum Disorder	67 (69)	43 (47)	38 (59)
Cognitive/developmental delays [a]	49 (50)	64 (65)	14 (20)
Attention Deficit	31 (32)	22 (24)	14 (22)
Anxiety	11 (12)	21 (23)	16 (25)
Self-Injury	12 (13)	11 (12)	5 (8)
OCD	3 (3)	11 (12)	4 (6)
Tics	4 (4)	2 (2)	2 (3)
Depression	1 (1)	5 (5)	4 (6)
Bipolar	1 (1)	1 (1)	1 (2)
Schizophrenia	1 (1)	0	1 (2)
Eating Disorders	2 (2)	0	2 (3)

[a] Parent reported concerned about cognitive development on developmental history form.

Table 2

Quartiles and Range of Intellectual Functioning, Language, and Adaptive Functioning by Synaptopathies

Measure	PMS (N = 98)			TSC (N = 98)			PHTS (N = 69)		
	n	Median [IQR]	Range	n	Median [IQR]	Range	n	Median [IQR]	Range
Intellectual Functioning									
Verbal IQ	97	16.35 [8, 43]	1-93	92	52.89 [43, 70]	4-119	64	80.50 [47, 104]	12-140
Non-Verbal IQ	97	24.49 [15, 43]	2-91	93	61.00 [43, 73]	4-115	66	76.00 [50, 96]	18-137
Full Scale IQ	97	19.90 [11, 41]	3-88	92	54.00 [40, 71]	4-106	64	77.00 [47, 99]	15-138
Adaptive Functioning [a]									
Communication	98	46.00 [40, 62]	26-91	87	65.00 [57, 75]	28-106	57	77.00 [61, 96]	29-129
Daily Living Skills	97	53.00 [40, 62]	25-89	89	65.00 [57, 76]	25-111	57	75.00 [63, 97]	41-125
Socialization	98	55.00 [48, 66]	34-101	86	68.00 [60, 83]	37-105	56	75.50 [59, 95]	24-127
Motor Skills	47	56.00 [51, 61]	25-81	38	68.50 [56, 75]	27-97	24	70.00 [63, 80]	54-104
Composite	98	49.00 [43, 62]	25-86	83	64.00 [56, 73]	26-103	55	76.00 [60, 90]	45-127
Language									
EVT-2	31	50.00 [30, 71]	20-92	69	73.00 [53, 86]	20-125	59	87.00 [70, 106]	20-139
PPVT-4	33	51.00 [25, 71]	19-82	75	74.00 [52, 86]	20-120	62	93.50 [54, 112]	20-144

Note. Interquartile range (IQR) [25th, 75th percentiles]

[a] Vineland Adaptive Behavior Scales-II domain and composite (ABC) standard scores. Motor Skills calculated for participants under 7 years old.

Table 3

Challenging Behavior Measures by Synaptopathies

Measure	PMS (N = 98)			TSC (N = 98)			PHTS (N = 69)		
	n	% Elevated	Mean (*SD*)	n	% Elevated	Mean (*SD*)	n	% Elevated	Mean (*SD*)
ABC subscale raw scores [a]									
Irritability	92	44	8.27 (8.50)	89	58	11.56 (10.30)	60	32	6.63 (7.65)
Lethargy	91	54	9.67 (9.00)	88	44	7.94 (8.17)	60	50	7.94 (8.27)
Stereotypy	92	61	5.10 (5.14)	89	47	4.35 (5.14)	60	52	3.92 (4.88)
Hyperactivity	91	63	18.82 (12.98)	89	53	15.07 (11.45)	60	30	10.27 (9.25)
Inappropriate Speech	92	41	1.91 (2.57)	89	53	2.84 (3.20)	60	60	2.75 (2.64)
CBCL scale T-scores [b]									
Attention problems				28	54	61.21 (7.90)	34	41	58.00 (8.25)
Externalizing				29	41	57.31 (12.44)	34	12	45.38 (11.80)
Internalizing				29	52	57.66 (10.45)	34	41	54.24 (13.87)
Anxiety				25	40	58.32 (6.81)	34	41	58.15 (8.42)

[a] ABC raw subscale scores above the mean of the U.S. special education community sample (by age and sex) are considered elevated/clinical risk.

[b] CBCL T-scores ≥ 60 meet the borderline clinical cut-off. CBCL scores are valid for participants with NVIQ ≥ 70. Scores from the PMS group are not presented because fewer than ten participants had valid NVIQs for administration. CBCL Internalizing includes Anxiety; Attention problems does not contribute to either Internalizing or Externalizing.

Table 4

Social-Communication, Repetitive Behavior and ASD Symptoms by Synaptopathies

Measure	PMS (N = 98)		TSC (N = 98)		PHTS (N = 69) [a]	
	n (%)	Mean (*SD*)	n (%)	Mean (*SD*)	n (%)	Mean (*SD*)
ADOS-2 [b]						
Module 1 ≥ 18 m MA	26 (27)	-	23 (23)	-	14 (20)	-
Module 2 ≥ 18 m MA	13 (13)	-	18 (18)	-	7 (10)	-
Module 3 ≥ 18 m MA	10 (10)	-	35 (36)	-	8 (12)	-
Module 4 ≥ 18 m MA	3 (3)	-	3 (3)	-	5 (7)	-
Calibrated Severity Score [c]	45 (46)	5.82 (2.63)	75 (77)	4.40 (3.24)	27 (39)	6.33 (2.08)
RBS-R						
Ritualistic/Sameness	91 (93)	4.42 (5.42)	90 (92)	6.76 (6.85)	60 (87)	5.55 (6.64)
Self-injurious Behavior	91 (93)	2.10 (2.56)	90 (92)	3.53 (4.63)	60 (87)	1.87 (2.73)
Stereotypic Behavior	91 (93)	5.21 (4.54)	88 (90)	5.76 (5.72)	60 (87)	5.73 (6.29)
Compulsive Behavior	91 (93)	1.58 (2.46)	89 (91)	2.47 (3.09)	59 (86)	2.63 (3.30)
Restricted Interests	91 (93)	2.63 (2.55)	90 (92)	2.96 (2.57)	60 (87)	2.90 (2.83)
Sum of factor scores	91 (93)	15.93 (13.26)	87 (89)	21.21 (18.01)	59 (86)	18.85 (17.46)
Total all RBS-R items	90 (92)	17.32 (15.46)	86 (88)	23.66 (20.19)	59 (86)	21.81 (19.92)
SRS-2						
Short 16-item total	86 (87)	27.73 (8.26)	87 (89)	25.31 (9.66)	60 (87)	24.82 (9.75)
T-score [d]	6 (6)	64.33 (8.85)	25 (26)	63.76 (10.97)	30 (43)	59.40 (14.59)

[a] ASD evaluations, including the ADOS-2, were performed only for PHTS participants for whom there was a history of or clinical impression of ASD. All participants in PMS and TSC were assessed for ASD.

[b] Nonverbal mental age (MA) based on nonverbal IQ or averaged Mullen Nonverbal Developmental Quotient for Fine Motor and Visual Reception.

[c] Sample size varies for Calibrated Severity Scores due to module administered, age maximums, and missing data at random.

[d] SRS-2 T-scores presented only for participants with NVIQ ≥ 70. Thirteen additional participants met NVIQ criteria, but nine (4 TSC, 5 PHTS) did not complete the SRS-2 and four (1 PMS, 3 PHTS) were missing items needed to calculate a T-score.

Table 5

ASD Classification by Measure and Condition

Measure	PMS (N = 98)	TSC (N = 98)	PHTS (N = 69)
Consensus Diagnosis, n (% Total) [a]	92 (94)	91 (93)	63 (91)
Diagnosis ASD, n (% Evaluated)	54 (59)	40 (44)	39 (62)
Certainty Rating, median [IQR]	5 [4, 5]	4 [4, 5]	5 [4, 5]
Diagnosis Non-ASD, n (% Evaluated)	38 (41)	51 (56)	24 (38)
Certainty Rating, median [IQR]	4 [3, 5]	5 [4, 5]	5 [4, 5]
ASD Assessment [b]			
History, n (% Total)	97 (99)	91 (93)	65 (94)
Autism, n (% Evaluated)	67 (69)	43 (47)	38 (59)
Non-autism, n (% Evaluated)	30 (31)	48 (53)	27 (41)
ADI-R Category [c], n (% Total)	49 (49)	64 (65)	40 (58)
Autism, n (% Evaluated)	37 (76)	30 (47)	31 (78)
Non-autism, n (% Evaluated)	12 (24)	34 (53)	9 (22)
Sensitivity	.88 (23/26)	.83 (25/30)	.90 (27/30)
Specificity	.41 (9/22)	.88 (29/33)	.60 (6/10)
ADOS-2 [c], n (% Total)	52 (53)	79 (81)	46 (67)
Autism, n (% Evaluated)	29 (56)	22 (28)	25 (54)
ASD, n (% Evaluated)	5 (10)	13 (16)	8 (18)
Non-autism, n (% Evaluated)	18 (34)	44 (56)	13 (28)
Sensitivity [d]	.96 (26/27)	.94 (30/32)	.91 (31/34)
Specificity	.65 (15/23)	.91 (41/45)	.83 (10/12)
ADI-R + ADOS-2 [e], n (% Total)	48 (48)	62 (63)	30 (43)
Sensitivity	.85 (22/26)	.77 (23/30)	.80 (24/30)
Specificity	.77 (17/22)	.97 (1/32)	.50 (5/10)

[a] Consensus diagnosis reflects clinician's impression based on all available data, captured on an ASD Consensus diagnosis form; participants missing this form are assumed to be missing at random. Two participants receiving a consensus diagnosis (PMS non-ASD, PHTS with ASD) had partial data, specifically, were missing autism clinical certainty rating on this form.

[b] ASD evaluations were performed only for PHTS participants for whom there was a history of or clinical impression of ASD. All participants in PMS and TSC were assessed for ASD.

[c] Valid administrations (mental age ≥ 18-months) were based on nonverbal IQ or estimates described in methods.

Note: Sensitivity (true positive divided by the sum of true positive and false negative) and Specificity (true negative divided by the sum of true negative and false positive) of ADI-R and ADOS-2 are defined relative to consensus diagnosis.

[d] Classifications of autism and ASD were combined for sensitivity & specificity calculations on the ADOS-2.

[e] Data compares positive autism/spectrum classifications on both the ADI-R and ADOS-2 classification to clinical consensus diagnosis