EXPERT REVIEW OF NEUROTHERAPEUTICS
https://doi.org/10.1080/14737175.2026.2655767

Taylor & Francis
by informa

CLINICAL TRIAL PROTOCOL

Rationale and study design for the first precision medicine randomized placebo-controlled trial in the 16p11.2 deletion syndrome

Hannah M. Rea[a,b], Julia T. Mattson[b,c], Vardan Arutiunian[d], Alyssa Verdes[e,f], Suvekcha Bhattachan[e,f], Megha Santhosh[d], Abigail Garcia[b], Anika Ajwani[b], Meaghan Bowen[g], Susan Faja[g], Ellen Hanson[g], Katherine Pawlowski[g], Robin P. Kochel[h*], Lisa Prock[g], Jeremy Veenstra-VanderWeele[e,f,i*], Karen Walton-Bowen[j,k], Adam Vogel[l,m], Angela Morgan[m,n], Sara Jane Webb[a,d*] and Paul P. Wang[j,o,p*]

[a]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA; [b]Institute of Human Development and Disability, University of Washington, Seattle, WA, USA; [c]Department of Pediatrics (Developmental Medicine), University of Washington, Seattle, WA, USA; [d]Center of Child Health, Behavior and Development, Seattle Children's Research Institute, Seattle, WA, USA; [e]Department of Child and Adolescent Psychiatry, New York State Psychiatric Institute, New York, NY, USA; [f]Center for Autism and the Developing Brain, New York Presbyterian Hospital, White Plains, NY, USA; [g]Division of Developmental Medicine, Department of Pediatrics, Boston Children's Hospital, Boston, MA, USA; [h]Division of Pediatrics-Psychology, Baylor College of Medicine, Houston, TX, USA; [i]Department of Psychiatry, Columbia University, New York, NY, USA; [j]Clinical Research Associates, LLC, New York, NY, USA; [k]Allos Pharma, Inc., Cambridge, MA, USA; [l]Redenlab Inc., Melbourne, Victoria, Australia; [m]Audiology and Speech Pathology, The University of Melbourne, Melbourne, Victoria, Australia; [n]Speech and Language, Murdoch Children's Research Institute, Parkville, Victoria, Australia; [o]Simons Foundation, New York, NY, USA; [p]Department of Pediatrics, Yale School of Medicine, New Haven, CT, USA

ABSTRACT

Chromosome 16p11.2 deletion syndrome is a genetic syndrome that includes difficulties in speech, language, and motor coordination. Arbaclofen, a selective GABA-B receptor agonist, has improved motor functioning and memory in mouse models. Prior clinical trials of arbaclofen in fragile X syndrome and autism spectrum disorder suggested benefit for social communication. L16hthouse (NCT04271332) is a multi-site, double-blind, randomized, placebo-controlled phase 2 trial to evaluate safety, efficacy, and tolerability of arbaclofen compared in 60 youths with 16p11.2 deletion syndrome (5 to 17:11 years) randomized on a 1:1 ratio. Primary outcomes included speech articulation, measured by the Goldman Fristoe Test of Articulation 3 (GFTA-3). Secondary outcomes included objective dysarthria indices, memory, motor control, and cognitive function, assessed with both standardized clinical measures and novel, computer-based assessments with automated scoring. Exploratory outcomes included attention, autism traits, and electrophysiological responses. L16hthouse is the first randomized trial in 16p11.2 deletion syndrome and uses an array of novel outcome measures to assess potential benefit in this population. In addition to providing potential insights about the safety, efficacy, and tolerability of arbaclofen, L16hthouse will provide an initial assessment of how these developmental outcome measures perform in a clinical trial across a broad age range.

Clinical trial registration number: NCT04271332; 2020-02-13.

PLAIN LANGUAGE SUMMARY

The methods are described for a multi-site, double-blind, randomized controlled Phase 2 trial to evaluate the safety, efficacy, and tolerability of arbaclofen for youth with 16p11.2 deletion syndrome. The primary outcome measure was speech articulation, a key difficulty for youth with 16p11.2 deletion syndrome. The outcomes of this trial, combined with the parallel Canadian ARBA and European AIMS-CT-01 trials, will contribute to the evidence base of arbaclofen as a treatment for neurological and psychiatric conditions.

ARTICLE HISTORY
Received 29 October 2025
Accepted 1 April 2026

KEYWORDS
16p11.2 deletion syndrome; arbaclofen; randomized controlled trial; speech; language; youth

1. Background & rationale

Deletions of a ~ 600 kilobase region between breakpoints 4 and 5 (BP4-BP5) on chromosome 16p11.2 are among the most frequently identified genetic abnormalities associated with autism, occurring in about 1% of individuals with a diagnosis of autism and about 0.6% of individuals with intellectual disability [1–4]. Approximately 3 in 10,000 people have been identified to have 16p11.2 deletion syndrome, with 60–76% of those cases being *de novo* mutations [5]. Among individuals with 16p11.2 deletion syndrome, common areas of difficulty include speech, language, memory, and motor coordination [1], with speech apraxia reported in a large majority of children [6]. In addition, it is estimated that approximately 22% of individuals with the deletion meet criteria for autism, 29% meet criteria for attention deficit hyperactivity disorder (ADHD), and about 30% meet criteria for intellectual disability [7]. There are no approved medication treatments for 16p11.2 deletion syndrome at this time.

CONTACT Sara Jane Webb sara.webb@seattlechildrens.org Center for Child Health, Behavior and Development, Seattle Children's Research Institute, 1920 Terry Ave., Seattle, WA 98101, USA; Paul P. Wang pwang@simonsfoundation.org Simons Foundation, 160 Fifth Avenue, 7th Floor, New York, NY 10010, USA
*Senior authors.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/14737175.2026.2655767

Article highlights

- Deletions of a ~ 600 kilobase region between breakpoints 4 and 5 (BP4-BP5) on chromosome 16p11.2 are among the most frequently identified genetic abnormalities associated with autism and intellectual disability.
- It is also estimated that ~3 in 10,000 people have been identified to have 16p11.2 deletion syndrome, with 60–76% of those cases being *de novo* mutations.
- The primary objectives of this study (the L16hthouse study) were to examine the safety, tolerability, and efficacy of arbaclofen among youth with 16p11.2 deletion with the primary outcomes for the trial being speech articulation. Secondary outcomes included memory, motor control, and cognitive function.
- To the authors' knowledge, this study represents the first-time speech articulation has been a primary endpoint in a precision medicine trial for a genetically defined neurodevelopmental condition.
- The current study has the potential to expand our understanding of the wider impacts of arbaclofen, to inform future research and therapeutic approaches, as well as to strengthen our understanding of how to measure potential treatment effects in RCTs in 16p11.2 deletion syndrome and other neurodevelopmental conditions.
- The L16hthouse study also represents a valuable model for designing RCTs using a collaborative, family-centered study design.

Early research in mouse models suggested that increased metabotropic glutamate receptor subtype 5 (mGluR5) signaling may contribute to the pathophysiology of 16p11.2 deletion [8]. As agonists of presynaptic gamma-aminobutyric acid (GABA-B) reduce mGluR activation and glutamate release, investigators tested whether mouse models of the syndrome might show benefit with the GABA-B receptor agonist, arbaclofen [9]. In findings replicated across multiple laboratories, these studies observed that arbaclofen rescued novel object recognition in two mouse models of the syndrome, as well as rescued the sequencing of motor behavior in an open field in one model [10]. Thus, arbaclofen appeared to show promise across 2 key symptom domains: memory and motor performance.

2. Previous arbaclofen trials

Arbaclofen is the active enantiomer of racemic baclofen, which is approved by the United States Food and Drug Administration (FDA) for the treatment of spasticity [9,11,12]. Arbaclofen has been investigated in clinical trials aimed at treating neurological and psychiatric conditions, including fragile X syndrome (FXS), autism spectrum disorder (ASD), multiple sclerosis (MS), gastroesophageal reflux disease (GERD), trigeminal neuralgia, and spasticity from spinal cord injury [13–21].

In a Phase 2 trial (2008–2010), arbaclofen was tested in a randomized, double-blind, placebo-controlled crossover study with individuals with FXS. The results showed improvements in social avoidance, measured using the Aberrant Behavior Checklist (ABC) [22] and the Vineland Adaptive Behavior Scale, Second Edition [23], as well as in challenging behaviors, measured using the visual analog scale problem behavior ratings [24]. Although subsequent Phase 3 studies did not achieve statistical significance on the primary clinical outcome assessment (ABC_{FX}-SA), clinically meaningful improvement were reported by caregivers and physicians in approximately half of the study subjects treated with arbaclofen in these trials [25]. A clear dose response was documented for subjects aged 5–11 years on behavioral and global outcome assessments [13]. The magnitude of improvement at the highest dose exceeded the clinically meaningful threshold for all three ABC_{FX} endpoints (ABC_{FX}-SA $p = 0.09$. ABC_{FX}-L/SW $p = 0.32$, and ABC_{FX}-I $p = 0.03$) where this threshold has been defined [26]. Furthermore, the highest doses were sufficiently well tolerated that it was considered safe to increase the maximum dosages for participants in this study by 33–50%.

Arbaclofen has also been assessed in three randomized, controlled trials (RCTs) in autism. A comparison of measures is presented in Supplementary Table S1. Phase 2 studies have been conducted by Seaside Therapeutics across 25 sites in the United States (NCT01288716, 2011–2012) [20], by the Autism Innovative Medicine Studies-2-Trials consortium (AIMS-2-Trials) in 7 sites across Europe (trial AIMS-CT-01, NCT03682978, 2019–2023) [18], and by the Province of Ontario Neurodevelopmental Network (POND) across 4 sites in Canada (trial ARBA, NCT03887676, 2019–2022). Veenstra-VanderWeele and colleagues reported modest improvements on the clinician-rated Clinical Global Impression of Severity and in a post-hoc analysis of social functioning on the Vineland Adaptive Behavior Scales [20]. In a sample of 122 autistic 5- to 17-year-olds, the AIMS-2-TRIALS-CT1 study found good adherence to protocol, and favorable safety. While they did not find significant differences on the primary or key secondary outcome measures, they did find significant improvements in the arbaclofen group compared to the control group in autism traits (measured using the AIM, the ABC, and the SRS-2) and quality of life. The authors also noted that their results were impacted by sample size and the interruption of data collection during COVID, and therefore should be considered in the larger context of research on arbaclofen [27]. ARBA and AIMS-CT-01 pre-registered a combined analysis, with the potential to address the sample size issue. Additional research has shown that single doses of arbaclofen may modulate EEG biomarkers of sensory processing, such as visual contrast perception and auditory habituation, in individuals with ASD [28,29]. In sum, recent research offers promising benefits for arbaclofen in youth with neurodevelopmental disorders.

The clinical trials of arbaclofen to date suggest that it is generally safe and well-tolerated and may have the potential to improve social and behavioral difficulties [16,20,24]. However, the overall clinical efficacy of arbaclofen remains unproven. The combination of promising findings in the mouse model and the well-defined safety and tolerability profile motivated a phase 2 RCT of arbaclofen in 16p11.2 deletion syndrome. The current trial was designed to largely overlap with existing trials to allow for comparisons between studies. Similar age ranges (5 to 11:11 year olds and 12 to 17:11 year olds) and dosing schedules were used (see section 2.2 Interventions), as were many screening and outcome measures (see section 2.4). While there were many similar assessments, primary and secondary outcome measures were identified in response to caregivers' primary concerns and characteristics of 16p11.2 deletion syndrome. For example,

the primary outcome of the current trial was speech motor difficulties (see below), rather than social skills, given caregiver focus on remediating speech challenges in this population. Similarly, the Differential Abilities Scale, Second Edition (DAS-II) was used in place of the Wechsler Scales as the current trial included a wider range of cognitive and verbal functioning. Other exploratory measures were added, as described below, that similarly captured salient concerns for youth with 16p11.2 deletion syndrome.

2.1. Objectives of the current study

The primary objectives of the L16hthouse study were to examine the safety, tolerability, and efficacy of arbaclofen among youth with 16p11.2 deletion. Speech motor difficulties were chosen as the primary objective because they are nearly universal among those with 16p11.2 deletion syndrome and families identify these difficulties as the primary concern. The primary outcome measure for efficacy was the Goldman Fristoe Test of Articulation 3 (GFTA-3) [30]. Secondary efficacy assessments included tests of memory and learning, fine motor control, body coordination, and cognitive functioning. Exploratory measures examined the impact of arbaclofen versus placebo on attention, verbal fluency, autism traits, motor function, and electrophysiological responses.

3. Methods

3.1. Study setting

The L16hthouse study was a multi-site double-blind, placebo-controlled Phase 2 trial where participants (up to $N = 60$) were randomized to arbaclofen or placebo at a 1:1 ratio. The trial was registered in ClinicalTrials.gov (NCT04271332).

Ethical approval for the trial was obtained from anonymized. Written informed consent was given by the participant's parent/caregiver/legally authorized representative (LAR) and, when applicable, the participant, in line with local laws and regulations. Assent from the participant was obtained, as appropriate. An independent Data Monitoring Committee (DMC) comprised of expert clinicians, and a statistician provided oversight of the trial conduct and monitored safety, efficacy, and dose escalations.

3.2. Eligibility criteria

(1) Diagnosis of 16p11.2 (BP4-BP5) deletion, with no additional known genetic disorders. Diagnosis was confirmed via review of medical records that were provided by families. The deletion could be de novo or inherited.
(2) Male or female participants 5;0 to 17;11 months of age, inclusive, at Screening.
(3) Neurodevelopmental disability requiring current educational or other therapeutic support. Any educational, behavioral, and/or other therapeutic interventions must have been continuous during the 2 months prior to screening, including stability of schooling and therapeutic services during the study.
(4) Participants with a history of seizure disorder must have been seizure-free on a stable antiepileptic therapy regimen for 6 months (if receiving anti-epileptic drugs [AEDs]) or must have been seizure-free for 3 years if not currently receiving AEDs. Individuals taking the anti-seizure medications vigabatrin and tiagabine, and many anxiolytics that also act as anti-seizure medications (see criteria 7) were excluded. If currently taking AEDs that are typically monitored by serum concentration (e.g. valproic acid, carbamazepine, or phenytoin), serum AED concentrations must be tested and confirmed to be within the therapeutic range at screening. Dosage of AEDs may be adjusted, and serum levels rechecked. If serum levels are confirmed to be within the therapeutic range after stable dosing for 30 days, this criterion may be considered met.
(5) No additional medical, neurologic, or psychiatric condition that might confound performance on assessment measures (e.g. significant impairment in hearing or vision, severe motor impairment from cerebral palsy, birth injury, or other injury, or cleft lip or palate), might interfere with the conduct of the study or interpretation of study results or endanger the participant's well-being (e.g. substance use disorder, impairment of renal function, evidence or history of malignancy or any significant hematological, endocrine, cardiovascular, respiratory, hepatic, or gastrointestinal disease).
(6) All medication regimens and dietary interventions must be stable for 30 days prior to screening and no new pharmacologic or non-pharmacologic interventions will be introduced during the study.
(7) Participants have not been treated with racemic baclofen or any investigational drug within 30 days prior to screening and are not currently treated with antipsychotic medications, more than 2 psychoactive medications (not including sleep aids that are used as needed), anxiolytic medications (e.g. buspirone, beta-blockers, benzodiazepines), vigabatrin, tiagabine, or riluzole.
(8) No known hypersensitivity to racemic baclofen.
(9) Prior to the conduct of any study-specific procedures, the participant must provide assent to participate in the study (if developmentally appropriate), and the parent/caregiver/LAR must provide written informed consent.
(10) The participant's parent/caregiver/LAR must be able to speak and understand English sufficiently to understand the nature of the study and to allow for the completion of all study assessments. The parent/caregiver/LAR should be capable of providing reliable information about the participant's condition, agree to oversee the administration of the study drug, and accompany the participant to all clinic visits. The same parent/caregiver/LAR should accompany the participant to each visit.
(11) Negative pregnancy test for females ≥9 years of age and, if sexually active, agreement to consistently use an accepted form of contraception.

3.3. Interventions

A study schematic is depicted in Figure 1. All participants attended a Screening Visit to determine eligibility. Eligible participants completed Visit 1 within 30 days of screening to collect baseline assessments and were randomized into either the arbaclofen or placebo, stratified by age group (5 to 11:11-year-olds or 12 to 17-year-olds). Arbaclofen and placebo were formulated as orally-disintegrating tablets of essentially identical shape, mass, color, smell, taste, and packaged in blister strips. Following Visit 1, medication titration to establish dosing lasted up to 35 days. On day 36 (+/- 3 days), participants attended Visit 2. Participants then maintained stable dosing for 77 days, followed by Visit 3 (day 113 +/- 3 days). Medication withdrawal occurred for 21 days and was followed by Visit 4 (day 134 +/-3 days). Thus, the double-blind portion of the study lasted between 164 and 193 days. Participants who completed the double-blind portion of the trial, adequately followed the protocol, and did not have medical contraindications were eligible to enroll in an Open-Label Extension. Participants, caregivers, and all study staff were blinded, except for a pharmacist at each site.

The dosing schedule was consistent with the European and Canadian autism trials [18]. All participants were titrated up, depending on tolerability. For 5 to 11-year-olds, participants began at a dose of 5 mg one time per day, and the maximum allowed dose was 15 mg three times per day. For 12 to 17-year-olds, the starting regimen was 5 mg twice per day, with the maximum-allowed dose at 20 mg three times per day. If a dose increase was not tolerated, the clinician would titrate the participant down to the previous dose and would not increase it again. At each visit, leftover pills were accounted for, and new pill packages were administered.

3.4. Outcomes

The schedule of assessments is presented in Table 1.

3.4.1. Eligibility measures

Following informed consent/assent, inclusion/exclusion criteria were reviewed, and medical, surgical, diagnostic, and treatment history (behavioral and pharmacological) were assessed via interview with the caregiver. Caregivers were also asked about changes in each domain at the start of subsequent visits.

3.4.2. Characterization measures

At screening, participants' autism traits were assessed using the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) [31], the Social Communication Questionnaire (SCQ) [32], and the Diagnostic and Statistical Manual, 5th Edition checklist [33].

3.4.3. Primary measures

3.4.3.1. Safety. Adverse Events were medical occurrences during the course of the study that were reported to the principal investigator (PI) by the parents and/or participants during phone call monitoring, face-to-face visits, medical examination, or laboratory testing. The nature of the event, onset, duration, and severity were recorded, and the PIs used standard definitions to determine the severity of the event (serious versus not serious, mild, moderate, severe) and relation to the drug (definitely related, probably related, possibly related, unlikely to be related, or not related). Clinically significant adverse events were accompanied by clinical symptoms, led to a change in study drug usage, and required a change in concomitant therapy. Per International Council for Harmonization (ICH) criteria, serious adverse events were those that result in death, were life-threatening, and/or resulted in inpatient hospitalization/prolonged hospitalization [34]. Adverse events were coded using the Medical Dictionary for Regulatory Activities (MedDRA) and summarized by system organ class and preferred term [34].

Safety of the medication was also tracked through physical examinations and blood tests. Full physical examinations conducted at Screening and Visit 3 included assessment of

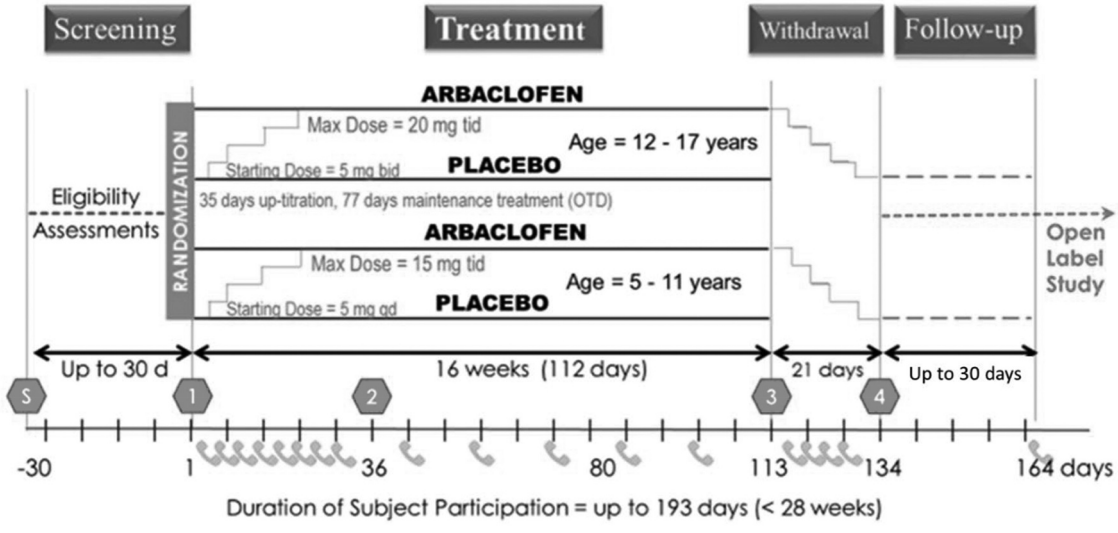

Figure 1. Dosing titration schedule, stratified by age.

Table 1. Study measures at each time point.

	Eligibility		Treatment period		Closeout
	Screening	V1	V2 (Day 36)	V3 (Day 113)	V4 (Day 134)
Eligibility					
Informed consent & assent	X				
Inclusion/exclusion criteria	X	X	X	X	X
Medical/Surgical History	X	X	X	X	X
Intervention History	X	X	X	X	X
Concomitant medications*	X	X	X	X	X
Characterization					
Autism characterization (DSM-5, SCQ, ADOS-2)	X				
Outcomes: Primary					
GFTA-3, speech battery		X		X	
Outcomes: Secondary					
BOT-2		X		X	
DAS-II	X			X	
WRAML-2		X		X	
Outcomes: Exploratory					
Motor testing with video recording		X	X	X	
CKAT		X	X	X	
EEG		X	X	X	
Conners' CPT-2		X	X	X	
D-KEFS		X	X	X	
AIM	X			X	
RBS-R	X			X	
Motor speech testing with audio recording		X		X	
Safety					
Adverse Events*		X	X	X	X
Risk assessment/Suicidality	X	X	X	X	X
Physical examination	X	X	X	X	X
Hematology	X	X		X	
Urinalysis	X	X		X	
Adherence					
Dose Monitoring Form*		X	X	X	X

Abbreviations: DSM-5 = Diagnostic and Statistical Manual 5th Edition; SCQ = Social Communication Questionnaire; ADOS-2 = Autism Diagnostic Observation Schedule 2nd Edition; AIM = Autism Impact Measure; RBS-R = Repetitive Behavior Scale—Revised; DAS-II = Differential Ability Scale 2nd Edition; GFTA = Goldman Fristoe Test of Articulation 3rd Edition; BOT-2 = Bruininks-Oseretsky Test 2nd Edition; CPT-2 = Conners' Continuous Performance Test; D-KEFS = Delis Kaplan Executive Function System; WRAML-2 = Wide Range. *Also monitored during phone calls.

cardiovascular, respiratory, gastrointestinal, and neurological systems, height and weight tracking, and Tanner Stage pubertal development review. Brief physical examinations conducted at Visit 1, Visit 2, and Visit 4 included abdominal palpitations to assess for liver enlargement, assessments of tone and posture, including orofacial tone, head control, and difficulty with other motor actions [35]. Vital signs, including temperature, heart rate, respiratory rate, and blood pressure were also assessed. Urine tests were conducted to screen for drug usage and, as appropriate, pregnancy. Suicidal ideation was assessed by a physician or clinical psychologist using a semi-structured interview with the parent and/or participant, as appropriate.

3.4.3.2. Speech articulation. While only a minority of persons with the 16p11.2 deletion have autism diagnoses, the vast majority have speech articulatory difficulties motivating the decision to focus on articulation as the primary endpoint [6]. The Goldman Fristoe Test of Articulation, 3rd Edition (GFTA-3) [30] Sound in Words and Sounds in Sentences are standardized tests of articulation of single words and words in connected speech, respectively, for ages 2 years 0 months to 21 years 11 months. Standard Scores are produced with a mean of 100 and a standard deviation of 15. Intelligibility during connected speech is also rated on a 4-point Likert scale for each sentence. The total of all sentence ratings is calculated, as is the overall intelligibility rating percentage. It was administered by a trained speech-language pathologist using the software platform by Redenlab Inc. at Visit 1 and Visit 3. The GFTA-3 and exploratory motor speech recordings (2.4.5.5 below) were made using a high-quality uni-directional cardiod microphone (xss) coupled with an external sound card (Roland Rubix 24) (Roland, Japan) sampled at 44.1 Hz and 16 bit quantization. As the GFTA-3 has not been used as an outcome measure in other clinical trials, its use is somewhat exploratory.

3.4.4. Secondary measures

3.4.4.1. Memory. The Wide Range Assessment of Memory and Learning, 2nd Edition (WRAML-2) [36] is a test of learning and memory for 5 to 90-year-olds. The 6 core subtests, 2 of the delay recall subtests, and 4 of the recognition subtests were administered. Together, the core subtests yield the Verbal Memory, Visual Memory, and Attention/Concentration indices that comprise the General Memory Index. Additionally, Verbal Recognition, Visual Recognition, and General Recognition Index Scores were calculated. As an experimental protocol, the Story Memory Delay Recognition and Picture Memory Delay Recognition subtests were re-administered following a 24-hour delay. Index scores are standard scores with a mean of 100 and a standard deviation of 15. The WRAML-2 was administered by a clinical psychologist during the Visit 1 and Visit 3.

3.4.4.2. Motor. The Bruininks-Oseretesky Test of Motor Proficiency, 2nd Edition (BOT-2) [37] is a measure of fine and gross motor abilities for 4 through 21-year-olds. The Fine Motor Precision and Balance subtests were administered. The Fine Motor Precision subtests require participants to complete activities that require precise finger and hand control (e.g. drawing, folding, cutting). The Balance subtests assess motor-control skills necessary for maintaining posture while standing, walking, or reaching.

3.4.4.3. Cognition. The Differential Ability Scales, 2nd Edition (DAS-II) [38], is an intelligence quotient (IQ) assessment for 2 years, 6 months through 17 years, 11 months. There are two overlapping age levels available, and both were implemented in this study: The Early Years battery is normed for children aged 2 years 6 months to 8 years 11 months, and the School Age battery is normed for children aged 7 years 0 months to 17 years 11 months. The DAS-II yields a Verbal Standard Score, Nonverbal Reasoning Standard Score, Spatial Standard Score, and General Conceptual Ability Standard Score, all of which have a mean of 100 with a standard deviation of 15. The DAS-II was administered by a clinical psychologist during the Screening Visit and Visit 3.

3.4.5. Exploratory measures

3.4.5.1. Electroencephalogram (EEG). Two sites (Seattle and Boston) provided EEG data collection using the EGI 128-channel system. EEG equipment description is in Supplemental Materials. In total, five EEG tasks were completed: The Autism Biomarkers Consortium for Clinical Trials (ABC-CT) Resting-State experiment consisted of 3 blocks of 1-minute-long videos of screensaver-like images [39]. From ARBA (NCT03887676), we included the Auditory Steady-State Response (ASSR) task in which participants were presented with 100 amplitude-modulated tones with the amplitude modulations of 10 Hz and 40 Hz while participants viewed a silent video of their choosing (e.g. My Little Pony, Pokémon). The ARBA Auditory Oddball experiment included the presentation of standard tones (1000 Hz) and deviant tones (2000 Hz) while participants watched a silent video. The AIMS-2-Trials (NCT03682978) Dynamic Social/Nonsocial Attention [18] paradigm consisted of the presentation of continuous videos with social (vignettes of women telling or singing nursery rhymes) and nonsocial (dynamic toys, such as a ball dropping down a chute) content (2 blocks of each). The ABC-CT Visual Evoked Potential paradigm [39] consisted of reversing black-and-white checkerboards. See additional information about EEG data collection in Supplementary Materials.

3.4.5.2. Attention. Conners' Continuous Performance Test (CPT) Kiddie CPT, 2nd Edition (Conners, 2015) [40], was administered to 5- to 7-year-olds, and the Conners' CPT, 3rd Edition (Conners, 2014) [41], was administered to participants 8 years and older. The CPT is a computerized, performance-based assessment of inattentiveness, impulsivity, sustained attention, and vigilance. Participants respond by pressing buttons in response to visual stimuli shown sequentially on a computer screen.

3.4.5.3. Autism traits. The Autism Impact Measure (AIM) [42] is a 41-item parent-report questionnaire designed to assess the severity and impact of autism traits on individuals across various domains, including social communication, behavioral symptoms, functioning, and family and community impact. Items are rated on 5-point Likert scales to measure frequency and impact.

The Repetitive Behavior Scale – Revised (RBS-R) [43] is a 43-item, parent-report questionnaire designed to assess the frequency and severity of repetitive and restricted behaviors in autistic individuals. Parents rate their child's behaviors on a 4-point Likert scale to yield T-scores for subdomains sensorimotor behaviors, restricted interests, self-injurious behaviors, compulsive behaviors, a need for sameness, and for total restricted, repetitive behaviors.

3.4.5.4. Motor. The Computer Vision Assessment of Motor Function [44] includes four tasks to assess developmental domains that are typically impacted in individuals with 16p11.2 deletion syndrome: social communication skills, gross motor skills, and oral motor functioning. The tasks employ a 5-minute face-to-face conversation to assess social communication, a test of imitation of gross motor movements, a gait test for walking coordination, and an evaluation of oral motor functioning during the GFTA-3. The participants were video recorded during each task, and the Center for Autism Research at the Children's Hospital of Philadelphia used machine learning to assess for subtle shifts in behavior.

Children's Kinematic Assessment Test (CKAT) [45] is a 12-to-15-minute tablet-based assessment for evaluating fine motor skills and visual-motor integration. Participants draw, trace, and connect dots on the tablet using an electronic stylus to yield scores related to speed and accuracy.

3.4.5.5. Motor speech testing with audio recording. The neurodevelopmental speech acoustic battery included five tasks to assess motor speech function: an automated task (days of the week), a voice quality and maximum phonation task (sustained vowel), fine motor timing (diadochokinesis – papapa, pataka), a picture description task and the children's test of non-word repetition (CNRep).Speech tasks were recorded as noted earlier [46].

3.5. Participant timeline

The project was primed for launch in early 2020 but was delayed because of the COVID-19 pandemic, with participant enrollment beginning in September 2022. The last participant enrolled in March 2024, and completion of all participant study activities occurred in February 2025.

3.6. Sample size

The planned sample size was $N = 60$ ($n = 30$ per treatment arm). The study was designed to provide preliminary assessments of safety, tolerability, and efficacy, so while statistically

underpowered, the proposed sample size is appropriate for a pilot/exploratory study and will provide useful information about feasibility and appropriateness of outcome measures [47–49].

3.7. Recruitment

Recruitment was restricted to North America and primarily conducted through anonymized Searchlight (https://www.simonssearchlight.org/), family meetings, and family Facebook groups. Data collection occurred anonymized

3.8. Statistical analysis

Analyses will be performed both with an Intent to Treat (ITT) and as Per Protocol (PP) approach. ITT analyses include all randomized participants who were administered at least 1 dose and have post-Visit 1 efficacy data. To be included in the PP sample, participants will fulfill ITT criteria and take at least 75% of the prescribed medication.

3.8.1. Safety

Safety summaries during the active double-blind experiment will be descriptive and include the number and percentage of participants experiencing adverse events, mean changes from baseline in health metrics as assessed during the physical examination, and the number and percentage of participants with clinically significant abnormal vital signs (i.e. temperature, heart rate, respiratory rate, and blood pressure) at each visit.

3.8.2. Efficacy

Efficacy will be analyzed by a difference between treatment arms at Visit 3 using Analysis of Covariance (ANCOVA), with Visit 3 score used as the dependent variable, treatment group as a factor, and baseline score as a covariate. Alternative analyses will include logistic regression and Chi-Square techniques, as appropriate for variable type. Hypothesis testing will be performed at the 5% level of significance for two-sided tests.

4. Methods: assignment of interventions

4.1. Sequence generation

The Interactive Web Response System (IWRS) was used for allocating and maintaining treatment masking of group assignment.

4.2. Allocation concealment mechanism

The investigators, study site staff, clinical research organization staff, and Medical Monitor did not have access to the treatment assignment, as the IWRS was responsible for group allocation.

4.3. Implementation

This was a double-blind, placebo-controlled (1:1 randomization), parallel group trial.

4.4. Masking

Unmasking of treatment allocation was allowed in case of emergency but was considered a protocol violation. Site investigators were to consult with the Medical Monitor prior to unmasking, unless a delay would jeopardize the safety and well-being of the participant. Medical monitors were to be informed of unmasking, but the participant's group assignment was not to be communicated to anyone unless necessary for ensuring safety and proper treatment. If unmasking occurred, the study drug would be discontinued and participation by that participant would end.

5. Methods: data collection, management, and analysis

5.1. Data collection methods

Site investigators were responsible for data collection, including the accuracy, completeness, and timeliness of the data. Site investigators were responsible for collecting and recording study data on source documents in a way that was accurate, clear, unambiguous, permanent, and capable of being audited. For example, ink was used, correction fluids or temporary attachments were forbidden, and photocopies/printouts of eCRFs were inadmissible. All sites were trained in October 2019 and were supported with ongoing training throughout the trial.

5.2. Data management

Source documents were kept in a secure, limited access area. Computer generated source documents were printed, signed, and dated by the investigator to become a permanent part of the subject's source documents.

Clinical report forms (CRFs) were entered in a database that conforms to FDA requirements for electronic data capture. Database electronic-CRFs were completed as soon as possible following the completion of a study visit, no more than 3 working days after the corresponding visit. Cross-check programs were used to ensure the completeness of electronic records. Error messages were generated to allow for the modification or verification of entered data. Queries were sent to the investigational site using an electronic data query system that included an automated audit trail of the corrections. The Sponsor contracted with a Clinical Research Organization (CRO) to monitor to review the source documents and electronic capture for completeness and accuracy; electronic-queries were generated to address inconsistencies, incomplete responses, or errors. The Principal Investigator certified data completeness and accuracy.

5.3. Methods

Demographic, baseline, medical history (coded using MedDRA terminology), concomitant medications (summarized by WHO Drug Dictionary class 4 names), safety, and efficacy data will be summarized and presented by treatment group and time point in summary tables. The diagnoses of speech disorders,

based on the GFTA-3 and oral-motor assessment conducted by the speech-language pathologist as part of the study and classified according to the Mayo Clinic dysarthria categorization (Articulation disorder, Phonological disorder, Stuttering, Dysarthria, Childhood Apraxia of Speech), will also be summarized descriptively (using frequencies and percentages) for the 2 treatment groups.

All statistical comparisons will be carried out as 2-sided tests. All efficacy analyses will be based on the intent to treat (ITT) population. The primary efficacy variable will also be analyzed in the PP population. Differences in primary and secondary efficacy variables from baseline to the end of 16 weeks of double-blind treatment in the arbaclofen group and placebo group will be assessed using Analysis of Covariance (ANCOVA) techniques for continuous variables and Chi-Square techniques for categorical variables, as appropriate. There will be no *p*-value adjustment for multiple comparisons in this Phase 2 exploratory study. Differences between the arbaclofen and placebo groups with respect to the change from baseline (last measurement prior to randomization) to the other study time points (e.g. Visit 2) will be summarized. The following information about treatment emergent AEs will be summarized: number and percentage of subjects, intensity, relationship to study drug, seriousness, and those that resulted in discontinuation of the study. Clinical laboratory variables, vital signs, physical examination findings, and suicidality assessments will be presented as change from baseline to each scheduled assessment, summarized by group. No imputation will be done for missing data in safety variables. A sensitivity analysis will be performed using multiple imputation to impute missing data for the primary efficacy analysis, but imputation of individual efficacy results will be implemented. The effects of site and age will be evaluated.

5.4. Methods: monitoring

5.4.1. Data monitoring

The following steps were taken to ensure the accuracy, consistency, completeness, and reliability of the data by the CRO: (1) Routine site monitoring; (2) eCRF review against source documents; (3) Data management quality control checks; (4) Statistical quality control checks; (5) Continuous data acquisition and cleaning; (6) Quality control of the final report.

5.4.2. Harms

See Section 2.4.3.1 for description of adverse events.

5.4.3. Auditing

A representative from the sponsor conducted periodic audits of the clinical sites and study processes. This included auditing the clinical database, final report, and study data.

6. Ethics and dissemination

6.1. Ethics approval and consent to participate

This study protocol was reviewed and approved by the internal review boards of Western anonymized written, informed consent was given by the participant's parent/caregiver/LAR and, when applicable, the participant, in line with local laws and regulations. Assent from the participant was obtained, as appropriate. An independent Data Monitoring Committee (DMC) comprised of expert clinicians, and a statistician provided oversight of the trial conduct and monitored safety, efficacy, and dose escalations. All procedures performed were in accordance with the ethical standards of the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

6.2. Dissemination policy

The policy and plan for data analysis and publications for the current study are derived from policies and practices that were utilized in other multicenter academic studies and are in accordance with the principles and standards of scientific research and scholarship within the fields of biomedical research and scientific journalism. The goals of this policy are: (1) to provide for the timely, scholarly, and comprehensive reporting of the data in the scientific literature; (2) to provide for the assignment of authorship and data analytic opportunities to study investigators in a manner that is equitable and supports career development; and (3) to ensure that the analysis and reporting data are consistent with regulatory agency requirements.

The data analysis and publication strategy will be determined by the Study Management Group. All study publications must be submitted for review to the Study Management Group.

Investigators are required to maintain and retain records for a period of 2 years following the date a marketing application is approved for the study drug for the indication for which it is being investigated. If no application is being filed, they must maintain and retain records for up to 2 years after the investigation is discontinued, and the FDA is notified. Custody of records may be transferred to the sponsor, who will follow retention policies.

7. Conclusion

The L16hthouse study was a double-blind randomized placebo-controlled trial designed to evaluate the safety, efficacy, and tolerability of arbaclofen in youth (5 to 17 years) with 16p11.2 deletion syndrome. The current study has the potential to advance the evidence base for a medication hypothesized to improve functioning for individuals with neurodevelopmental conditions more broadly, as arbaclofen is being evaluated in other populations, and was designed to overlap in methods with concurrent RCTs (e.g. POND, AIMS-2-CT1) to best contribute to the evidence base investigating the efficacy, safety, and tolerability of arbaclofen [18,20]. To the authors' knowledge, it also represents the first-time speech articulation has been a primary endpoint in a precision medicine trial for a genetically defined neurodevelopmental condition. In conjunction with those other clinical trials, positive outcomes related to arbaclofen could have meaningful implications for future trials on this and other

medications that target excitatory/inhibitory imbalances in autism.

L16hthouse is a novel and meaningful trial as it extends the research to a novel population and includes primary outcomes related to speech-motor difficulties, based upon the philosophy of patient-centered outcome research. In a survey of over 200 families of youth with 16p11.2 deletion syndrome, 60% of families indicated speech or language use as their primary concern. Further, 60% of caregivers also reported concerns related to learning, memory, and cognition, and 30% reported motor issues. The L16hthouse study was thus designed to integrate both standardized tasks in these target domains, as well as novel, objective, performance-based experimental measures of various facets of motor coordination, to elucidate other potential benefits of arbaclofen in response to families' observations/concerns. The current study has the potential to expand our understanding of the wider impacts of arbaclofen, to inform future research and therapeutic approaches, as well as to strengthen our understanding of how to measure potential treatment effects in RCTs in 16p11.2 deletion syndrome and other neurodevelopmental conditions. The L16hthouse study also provides a model for designing RCTs using a collaborative, family-centered study design.

Ethics approval and consent to participate

This study protocol was reviewed and approved by the internal review boards of WCG anonymized Written, informed consent was given by the participant's parent/caregiver/LAR and, when applicable, the participant, in line with local laws and regulations. Assent from the participant was obtained, as appropriate. An independent Data Monitoring Committee (DMC) comprised of expert clinicians and a statistician provided oversight of the trial conduct and monitored safety, efficacy, and dose escalations. All procedures performed were in accordance with the ethical standards of the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Data availability statement

Data from this study are now owned by Allos Pharmaceuticals, who are committed to open sharing of the data through collaborative projects. The data will be shared on clinicaltrials.gov.

SPIRIT 2025 guidelines

The Standard Protocol Items: Recommendations for Interventional Trials standards were followed.

Acknowledgments

We wish to thank the families, parents, and children who participated in the L16hthouse study. We also want to thank all of the staff and personnel who contributed to the research.

Author contributions

CRediT: **Hannah M. Rea:** Investigation, Project administration, Writing – original draft, Writing – review & editing; **Julia T. Mattson:** Investigation, Project administration, Writing – original draft, Writing – review & editing; **Vardan Arutiunian:** Resources, Software, Writing – original draft, Writing – review & editing; **Alyssa Verdes:** Data curation, Investigation, Project administration, Writing – original draft, Writing – review & editing; **Suvekcha Bhattachan:** Data curation, Investigation, Project administration, Writing – original draft, Writing – review & editing; **Megha Santhosh:** Data curation, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – review & editing; **Abigail Garcia:** Data curation, Investigation, Writing – review & editing; **Anika Ajwani:** Data curation, Investigation, Writing – review & editing; **Meaghan Bowen:** Data curation, Investigation, Project administration, Resources, Writing – review & editing; **Susan Faja:** Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review & editing; **Ellen Hanson:** Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review & editing; **Katherine Pawlowski:** Data curation, Investigation, Project administration, Resources, Supervision, Writing – review & editing; **Robin P. Kochel:** Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing; **Lisa Prock:** Data curation, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – review & editing; **Jeremy Veenstra-VanderWeele:** Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – review & editing; **Karen Walton-Bowen:** Conceptualization, Data curation, Funding acquisition, Methodology, Resources, Writing – review & editing; **Adam Vogel:** Conceptualization, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing; **Angela Morgan:** Conceptualization, Methodology, Project administration, Resources, Supervision, Writing – review & editing; **Sara Jane Webb:** Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – review & editing; **Paul P. Wang:** Conceptualization, Data curation, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review & editing.

Funding

This clinical trial was sponsored by Clinical Research Associates, LLC (CRA), a non-profit affiliate of the Simons Foundation, which provided funding to the authors involved in the study (V Arutiunian, A Verdes, S Bhattachan, M Santhosh, A Garcia, M Dyer, S Faja, E Hansen, K Pawlowski, R Kochel, L Prock, J Veenstra-VanderWeele, K Walton-Bowen, SJ Webb, and P Wang). H Rea was funded by [KL2TR002317]. Institutional support was provided to the University of Washington site through the National Institute of Child Health and Human Development (NICHD) funding ([IDDRC P50HD103524] awarded to A Ajwani, JT Mattson, SJ Webb, and H Rea).

Declaration of Interest

A Vogel works for Redenlab Inc, Australia, which created the software platform used to record intelligibility, one of the primary outcome measures. P Wang was employed by Clinical Research Associates, LLC and Seaside Therapeutics and is employed by the Simons Foundation. Karen-Walton Bowen previously worked at Seaside Therapeutics and Clinical Research Associates, LLC and currently works at Allos Pharma.

Reviewer disclosures

Peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Chung WK, Roberts TP, Sherr EH, et al. 16p11.2 deletion syndrome. Curr Opin Genet Dev. 2021;68:49–56. doi: 10.1016/j.gde.2021.01.011
   - **This manuscript provides important background information about 16p11.2 deletion syndrome.**

2. Weiss LA, Shen Y, Korn JM, et al. Association between microdeletion and microduplication at 16p11.2 and autism. N Engl J Med. 2008;358(7):667–675. doi: 10.1056/NEJMoa075974
3. Hanson E, Nasir RH, Fong A, et al. Cognitive and behavioral characterization of 16p11.2 deletion syndrome. J Dev Behavioral Pediatrics. 2010;31(8):649–657. doi: 10.1097/DBP.0b013e3181ea50ed
• **This manuscript provides important background information about characteristics associated with 16p11.2 deletion syndrome.**
4. Shinawi M, Liu P, Kang SHL, et al. Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. J Med Genet. 2010;47(5):332–341. doi: 10.1136/jmg.2009.073015 PubMed PMID: 19914906; PubMed Central PMCID: PMC3158566.
5. Rein B, Yan Z. 16p11.2 copy number variations and neurodevelopmental disorders. Trends Neurosci. 2020;43(11):886–901. doi: 10.1016/j.tins.2020.09.001
6. Mei C, Fedorenko E, Amor DJ, et al. Deep phenotyping of speech and language skills in individuals with 16p11.2 deletion. Eur J Hum Genet. 2018;26(5):676–686. doi: 10.1038/s41431-018-0102-x
7. Niarchou M, Chawner SJRA, Doherty JL, et al. Psychiatric disorders in children with 16p11.2 deletion and duplication. Transl Psychiatry. 2019;9(1):8. doi: 10.1038/s41398-018-0339-8
8. Tian D, Stoppel LJ, Heynen AJ, et al. Contribution of mGlur5 to pathophysiology in a mouse model of human chromosome 16p11.2 microdeletion. Nat Neurosci. 2015;18(2):182–184. doi: 10.1038/nn.3911
9. Isaacson JS, Hille B. GABAB-Mediated presynaptic inhibition of excitatory transmission and synaptic vesicle dynamics in cultured hippocampal neurons. Neuron. 1997;18(1):143–152. doi: 10.1016/S0896-6273(01)80053-2 PubMed PMID: 9010212.
10. Gundersen BB, O'Brien WT, Schaffler MD, et al. Towards preclinical validation of arbaclofen (R-baclofen) treatment for 16p11.2 deletion syndrome. bioRxiv. 2023:2023.05.01.538987. doi: 10.1101/2023.05.01.538987 PubMed PMID: 37745360; PubMed Central PMCID: PMC10515778.
11. Scanziani M, Capogna M, Gähwiler BH, et al. Presynaptic inhibition of miniature excitatory synaptic currents by baclofen and adenosine in the hippocampus. Neuron. 1992;9(5):919–927. doi: 10.1016/0896-6273(92)90244-8 PubMed PMID: 1358131.
12. Sohn JW, Lee D, Cho H, et al. Receptor-specific inhibition of GABAB-activated K+ currents by muscarinic and metabotropic glutamate receptors in immature rat hippocampus. J Physiol. 2007;580(2):411–422. doi: 10.1113/jphysiol.2006.125914
13. Berry-Kravis E, Hagerman R, Visootsak J, et al. Arbaclofen in fragile X syndrome: results of phase 3 trials. J Neurodevelop Disord. 2017;9(1):3. doi: 10.1186/s11689-016-9181-6
14. Erickson CA, Veenstra-Vanderweele JM, Melmed RD, et al. STX209 (arbaclofen) for autism spectrum disorders: an 8-week open-label study. J Autism Dev Disord. 2014;44(4):958–964. doi: 10.1007/s10803-013-1963-z
15. Mentch J, Spiegel A, Ricciardi C, et al. Gabaergic inhibition gates perceptual awareness during binocular rivalry. J Neurosci. 2019;39(42):8398–8407. doi: 10.1523/JNEUROSCI.0836-19.2019
16. Nance PW, Huff FJ, Martinez-Arizala A, et al. Efficacy and safety study of arbaclofen placarbil in patients with spasticity due to spinal cord injury. Spinal Cord. 2011;49(9):974–980. doi: 10.1038/sc.2011.43
17. Okuda DT, Kantor D, Jaros M, et al. Arbaclofen extended-release tablets for spasticity in multiple sclerosis: open-label extension study. Brain Comm. 2022;5(1):fcad026. doi: 10.1093/braincomms/fcad026
18. Parellada M, San José Cáceres A, Palmer M, et al. A phase II randomized, double-blind, placebo-controlled study of the efficacy, safety, and tolerability of arbaclofen administered for the treatment of social function in children and adolescents with autism spectrum disorders: study protocol for AIMS-2-TRIALS-CT1. Front Psychiatry. 2021;12:701729. doi: 10.3389/fpsyt.2021.701729 PubMed PMID: 34504446; PubMed Central PMCID: PMC8421761.
•• **The current methods were designed to be similar to the AIMS-2 TRIALS.**
19. Vakil NB, Huff JF, Bian A, et al. Arbaclofen placarbil in GERD: a randomized, double-blind, placebo-controlled study. Off J Am Coll Gastroenterol | ACG. 2011;106(8):1427. doi: 10.1038/ajg.2011.121
20. Veenstra-VanderWeele J, Cook EH, King BH, et al. Arbaclofen in children and adolescents with autism spectrum disorder: a randomized, controlled, phase 2 trial. Neuropsychopharmacol. 2017;42(7):1390–1398. doi: 10.1038/npp.2016.237
21. Fromm G, Terrence C. Comparison of L-baclofen and racernic baclofen in trigeminal neuralgia. Neurology. 1987;37(11):1725–1728.
22. Aman MG, Singh NN, Stewart AW, et al. Psychometric characteristics of the aberrant behavior checklist. Am J Ment Defic. 1985;89(5):492–502. PubMed PMID: 3158201.
23. Sparrow SS, Cicchetti DV, Balla DA. Vineland adaptive behavior scales (Vineland-II). 2nd ed. San Antonio (TX): Pearson; 2005.
24. Berry-Kravis EM, Hessl D, Rathmell B, et al. Effects of STX209 (arbaclofen) on neurobehavioral function in children and adults with fragile X syndrome: a randomized, controlled, phase 2 trial. Sci Transl Med. 2012;4(152). doi: 10.1126/scitranslmed.3004214
25. Sansone SM, Widaman KF, Hall SS, et al. Psychometric study of the aberrant behavior checklist in fragile X syndrome and implications for targeted treatment. J Autism Dev Disord. 2012;42(7):1377–1392. doi: 10.1007/s10803-011-1370-2
26. Berry-Kravis E, Hagerman R, Budimirovic D, et al. A randomized, controlled trial of ZYN002 cannabidiol transdermal gel in children and adolescents with fragile X syndrome (CONNECT-FX). J Neurodev Disord. 2022;14:1–15. doi: 10.1186/s11689-022-09466-6
27. Parellada M, Cáceres ASJ, Delorme R, et al. Efficacy, safety, and tolerability of arbaclofen in autistic children and adolescents, the AIMS-2-TRIALS-CT1: a randomized, double-blind, placebo-controlled phase ii trial. eClinicalmedicine. 2026;92. doi: 10.1016/j.eclinm.2026.103760 PubMed PMID: 41631164.
28. Huang Q, Velthuis H, Pereira AC, et al. Exploratory evidence for differences in GABAergic regulation of auditory processing in autism spectrum disorder. Transl Psychiatry. 2023;13(1):1–13. doi: 10.1038/s41398-023-02619-8
29. Huang Q, Ellis CL, Leo SM, et al. Retinal GABAergic alterations in adults with autism spectrum disorder. J Neurosci. 2024;44(14):e1218232024. doi: 10.1523/JNEUROSCI.1218-23.2024
30. Goldman R, Fristoe M. GFTA-3: Goldman Fristoe 3 test of articulation. PsychCorp; 2015.
31. Lord C, Rutter M, Dilavore PC, et al. ADOS-2: autism diagnostic observation schedule. Hogrefe; 2008.
32. Rutter M. The Social Communication Questionnaire. Western Psychological Services; 2003.
33. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Washington (DC): American Psychiatric Association; 2013.
34. International Council for Harmonisation (ICH). International council for harmonisation of technical requirements for pharmaceuticals for human use (ICH). Efficacy guidelines. [Internet]. 2013. Available from: https://www.ich.org/page/efficacy-guidelines
35. Robbins J, Klee T. Clinical assessment of oropharyngeal motor development in young children. J Speech Hear Disord. 1987;52(3):271–277. doi: 10.1044/jshd.5203.271 PubMed PMID: 3455449.
36. Sheslow D, Adams W. Wide range assessment of memory and learning (WRAML). Bloomington (MN): NCS Pearson; 2003.
37. Bruininks R, Bruininks B. Bruininks-Oseretsky test of motor proficiency. 2nd ed. Minneapolis (MN): NCS Pearson; 2005.
38. Elliot CD. Differential ability scales. 2nd ed. San Antonio (TX): The Psychological Corporation; 2007.
39. Webb SJ, Shic F, Murias M, et al. Biomarker acquisition and quality control for multi-site studies: the autism biomarkers consortium for clinical trials. Front Integr Neurosci [Internet]. 2020 [cited 2022 Oct 21];13. Available from: https://www.frontiersin.org/articles/10.3389/fnint.2019.00071
40. Conners C. Conners kiddie continuous performance test 2nd edition (Conners K-CPT 2) manual. 2015.

41. Conners CK. Conners continuous performance test 3rd edition (Conners CPT 3) & Connors continuous auditory test of attention (Conners CATA): technical manual. MHS; 2014.
42. Kanne SM, Mazurek MO, Sikora D, et al. The autism impact measure (AIM): initial development of a new tool for treatment outcome measurement. J Autism Dev Disord. 2014;44(1):168–179. doi: 10.1007/s10803-013-1862-3
43. Bodfish JW, Symons FJ, Parker DE , et al. Varieties of repetitive behavior in autism: comparisons to mental retardation. J Autism Dev Disord. 2000;30:237–243. doi: 10.1023/A:1005596502855
44. McDonald DQ, Sariyanidi E, Zampella CJ , et al. Predicting autism from head movement patterns during naturalistic social interactions. In: Proceedings of the 2023 7th International Conference on Medical and Health Informatics; Keyoto, Japan; 2023 May. p. 55–60. doi: 10.1145/3608298.3608309 PubMed PMID: 38699395; PubMed Central PMCID: PMC11064057.
45. Cunningham AC, Hill L, Mon-Williams M, et al. Using kinematic analyses to explore sensorimotor control impairments in children with 22q11.2 deletion syndrome. J Neurodev Disord. 2019;11(1):8. doi: 10.1186/s11689-019-9271-3 PubMed PMID: 31182009; PubMed Central PMCID: PMC6558818.
46. Gathercole SE, Willis CS, Baddeley AD, et al. The children's test of non-word repetition: a test of phonological working memory. Memory. 1994;2(2):103–127. doi: 10.1080/09658219408258940 PubMed PMID: 7584287.
47. McMenamin M, Berglind A, Wason JMS. Improving the analysis of composite endpoints in rare disease trials. Orphanet J Rare Dis. 2018;13(1):81. doi: 10.1186/s13023-018-0819-1 PubMed PMID: 29788976; PubMed Central PMCID: PMC5964664.
48. Mauguen A. Accounting for the rarity of the disease when designing clinical trials with a focus on pediatric cancers. Clin Trials. 2022;19(3):292–296. doi: 10.1177/17407745221080728 PubMed PMID: 35232288; PubMed Central PMCID: PMC9232871.
49. Partington G, Cro S, Mason A, et al. Design and analysis features used in small population and rare disease trials: a targeted review. J Clin Epidemiol. 2022;144:93–101. doi: 10.1016/j.jclinepi.2021.12.009 PubMed PMID: 34910979.