

Published in final edited form as:

J Dev Behav Pediatr. 2013 September ; 34(7): 508–522. doi:10.1097/DBP.0b013e31829d1f20.

Outcome Measures for Clinical Trials in Fragile X Syndrome

Elizabeth Berry-Kravis, MD, PhD¹, David Hessl, PhD^{2,3}, Leonard Abbeduto, PhD^{2,3}, Allan L. Reiss, MD⁴, Andrea Beckel-Mitchener, PhD⁵, Tiina K. Urv, PhD⁶, and Outcome Measures Working Groups*

¹Departments of Pediatrics, Neurological Sciences and Biochemistry Rush University Medical Center, Chicago, IL

²Department of Psychiatry and Behavioral Sciences, University of California, Davis School of Medicine, Sacramento, CA

³MIND Institute, University of California, Davis Medical Center, Sacramento, CA

⁴Center for Interdisciplinary Brain Sciences Research and Departments of Psychiatry and Behavioral Sciences, Radiology and Pediatrics Stanford University School of Medicine, Stanford, CA

⁵National Institute of Mental Health, National Institutes of Health, Bethesda, MD

⁶Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD

Abstract

Objective—Progress in basic neuroscience has led to identification of molecular targets for treatment in fragile X syndrome (FXS) and other neurodevelopmental disorders, however, there is a gap in translation to targeted therapies in humans. One major obstacle to the demonstration of efficacy in human trials has been the lack of generally accepted endpoints to assess improvement in function in individuals with FXS. To address this problem, the NIH convened a meeting of leading scientists and clinicians with the goal of identifying and standardizing outcome measures for use as potential endpoints in clinical trials in FXS.

Methods—Participants in the meeting included FXS experts, experts in the design and implementation of clinical trials and measure development, and representatives from advocacy groups, industry, and federal agencies.

Results—The group generated recommendations for optimal outcome measures in cognitive, behavioral, and biomarker/medical domains, including additional testing and validation of existing

Address correspondence to: Elizabeth Berry-Kravis, MD, PhD, Rush University Medical Center, 1725 West Harrison, Suite 718, Chicago, IL 60612, Phone: 312-942-4036, Fax: 312-942-4168, Elizabeth_m_berry-kravis@rush.edu.

***Outcome Measures Working Group Participants**

Behavior Work Group: Michael Aman, Katie Clapp, Scott Hall, David Hessl, Becky Kronk, Ave Lachiewicz, Robert Noll, Lawrence Scahill

Cognition Work Group: Leonard Abbeduto, Elizabeth Berry-Kravis, Heather Cody Hazlett, Stephen Hooper, Bryan King, Amy Lightbody, Richard Paylor, David Posey, Paul Wang, Becky Zorovic

Medical/Physical Work Group: Craig Erickson, Randi Hagerman, Walter Kaufmann, Eric Klann, Joseph Piven, Allan Reiss, Linmarie Sikich, Nicole Tartaglia, Michael Tranfaglia

Conflicts of Interest and Source of Funding: EBK has received funding to consult and conduct clinical trials in FXS from Novartis, Seaside Therapeutics, and Roche, and has received grant funding from the NIH, CDC, FRAXA Research Foundation, and the NFXF. DH has received funding to consult and conduct clinical trials in FXS from Novartis, Seaside Therapeutics, and Roche, and has received grant funding from the NIH, DOD, FRAXA Research Foundation, the John Merck Foundation and the NFXF. LA has received grant funding from the NIH and the NFXF. ALR has received consulting fees from Novartis and grant funding from the NIH. AB-M and TKU have nothing to disclose. The views expressed in this article are those of the authors and do not necessarily represent the views of the NIH or the United States Government.

measures, and development of new measures in areas of need. Although no one endpoint or set of endpoints could be identified that met all criteria as an optimal measure, recommendations are presented in this report.

Conclusion—The report is expected to guide the selection of measures in clinical trials and lead to the use of a more consistent battery of measures across trials. Further, this will help to direct research toward gaps in the development of validated FXS-specific outcome measures, and to assist with interpretation of clinical trial data by creating templates for measurement of treatment efficacy.

Keywords

fragile X syndrome; clinical trials; outcome measures; intellectual disability

Introduction

An important goal of research in the field of intellectual and developmental disabilities is to discover methods for ameliorating the impact of these disorders on health and quality of life of affected individuals and their caregivers. One such disorder is fragile X syndrome (FXS). FXS is an X-chromosome linked single-gene disorder and the most common known, heritable cause of cognitive and behavioral dysfunction in humans, with prevalence rates as high as 1 in 2,500 (1,2). FXS is caused by a trinucleotide repeat expansion mutation in the regulatory region of the Fragile X Mental Retardation 1 gene (*FMR1*) which results in hypermethylation and transcriptional silencing of the gene, and the resulting loss or reduction of expression of the gene product, Fragile X Mental Retardation Protein (FMRP). Insufficient expression of FMRP, an mRNA-binding translational regulator, has a broad array of effects on cellular signaling pathways and on synaptic plasticity, morphology and function, thereby leading to abnormalities in brain connectivity and neurodevelopmental processes (3–6). In addition to intellectual disability (ID), individuals with FXS often demonstrate clinical features that overlap with other frequently occurring psychiatric and developmental disorders including autism, ADHD, anxiety disorders, and language disorders (7).

A number of symptom-based pharmacological treatments for individuals with FXS have been employed, including stimulants for attention deficits, selective serotonin reuptake inhibitors (SSRIs) for anxiety and atypical neuroleptics for aggression (8,9). Unfortunately, there is limited available data regarding the efficacy of symptom-based pharmacological or behavioral treatments in FXS. Clinical trials of sufficient size and power to detect definitive treatment effects have not been conducted. Studies have primarily consisted of open-label pilot trials and small pilot placebo-controlled double-blind trials of symptom-based or mechanistically-targeted pharmacotherapy (Table 1). While these trials have provided valuable information about the potential efficacy of approved and investigational drugs in FXS, they have also demonstrated multiple shortcomings of currently available outcome measures. Further, clinical endpoints used to assess efficacy have differed across medication trials in FXS, making cross-study comparisons difficult (Table 1). Moreover, the adequacy of many of those clinical measures remains to be determined.

New knowledge derived from studies of translational models of FXS, as well as genetic, imaging, and neuropsychological investigations of people with FXS has opened the door to the development of disease-specific pharmacological treatments (10), with several drugs currently being tested in Phase 2 and 3 clinical trials (Table 1). Prompted by the extraordinary research progress, and in anticipation of ongoing acceleration of the clinical trial effort in humans with FXS, the National Institutes of Health convened a meeting of leading scientists and clinicians with the goal of identifying outcome measures for use as

potential endpoints in clinical trials in FXS. Participants in the meeting included FXS experts, along with experts in the design and implementation of clinical trials and measure development, and representatives from advocacy groups, the pharmaceutical industry, and various federal agencies. The overarching objective of this meeting was to facilitate discussion that would lead to the identification and/or development of a set of outcome measures that could be used to document efficacy in current and future clinical trials.

The work of the participants in the meeting was guided by three principles. First, although it is unlikely that a single battery of measures would be appropriate for all clinical trials, identification of a core set of widely applicable measures would facilitate comparability across different agents, research centers, and methodological approaches. Second, outcome measures that have been developed for symptom-based clinical trials in behaviorally defined disorders (e.g., autism, ADHD) might not be sufficiently sensitive or specific for disease-oriented interventions in FXS and thus, measures must be validated specifically for FXS. Third, the results from the meeting should generate a set of criteria for identifying appropriate outcome measures that could help inform the design of clinical trials, improve interpretation of clinical trial data by regulatory agencies such as the FDA, and help guide governmental and private sector funding entities on this topic.

Methods and Procedures

Working groups were created to identify outcome measures in three areas relevant to the broad phenotype of FXS: (1) Cognition; (2) Behavior/Emotion; and (3) Medical/Physical including Biomarkers (Figure 1). Each working group included individuals with clinical, research, and measurement expertise in developmental disorders. Each of the three working groups participated in a series of meetings, during which the process described in Figure 1 was carried out and all relevant published literature was evaluated. Following several months of virtual meetings, a 2-day face-to-face meeting was convened in Washington, D.C. The meeting included FDA input on the process, presentation of working group progress, intensive work within working groups to refine their list of optimal endpoints, and a final open discussion.

Results

Summaries of data analyses and discussion from each working group are presented in the following sections. It should be noted that although much of the activities of the working groups focused on parsing cognition, behavior and emotion, and biological systems into distinct subdomains, there was also the recognition of the possibility of developing a single composite measure for each domain – a composite that aggregated the core features of FXS within that domain. In this way, investigators might seek approval for a therapeutic designed to treat a cluster of related features associated with FXS rather than just specific symptoms. This single-composite approach is likely to have some practical challenges, such as how to identify features that will be included in the measure, as well as variation in symptom presentation across age and level of impairment. Nonetheless, such an approach is appealing in light of neurotherapeutics in development, which target the underlying neurobiology of FXS. A potential drawback of a composite measure is that it may fail to identify true clinical improvement in a single functional domain, as may occur when an intervention has a narrow, yet clinically significant effect, on only one of the aggregated subdomains. The single-composite approach was endorsed most enthusiastically by the Behavior and Emotion working group.

Cognitive Measures

Although use of a single measure of cognitive function is appealing, the working group was skeptical that this approach would be robust for clinical trials in FXS given the rather modest changes in cognition expected from various classes of drugs being tested. Focused assessments of core cognitive features seem better suited for clinical trials at this time.

Given this starting point, the working group was guided by the following:

1. The therapeutics currently being developed are designed to impact neural systems (e.g., the mGluR5 pathway) that are related to numerous subdomains of cognition as well as aspects of emotion and behavior. For cognition, clinical effects are likely to be observed in both FXS-specific features of cognition (e.g., high rates of perseverative language) and in features that are shared by other neurodevelopmental disorders (e.g., limited short-term memory).
2. Improvements in cognition can be measured in a variety of ways. On the one hand, improvements can be reflected in the “products” of learning, for example an increase in the number of words that an individual can produce. On the other hand, improvements can be reflected in the “processes” underlying learning, as in the case of an accelerated learning rate (11). Process measures may be more valuable for evaluating clinically meaningful effects because they reflect improvements in the capacity to learn and adapt, which should have significant effects on daily functioning and quality of life. Process measures may also be more sensitive to drug effects than “product” measures, which are essentially the result of a learning process and therefore “accumulate” to measurable levels gradually.
3. The scope and nature of changes in cognition vary with age. Thus, different measures will be needed for participants of different ages.
4. Although it is useful to distinguish different areas of functioning, such as cognition and behavior, these domains are interrelated and measurement strategies should be designed accordingly. The working group found value in partitioning cognition into the subdomains of language, memory and learning, executive functioning, social cognition, and academic achievement; again, recognizing that these areas are inter-related. A list of promising outcome measures for the cognitive domain can be found in Table 2.

Language—The linguistic phenotype of FXS is characterized by language delays relative to chronological age expectations (12–17). Atypical linguistic behaviors are also common, including high rates of verbal perseveration (18,19), and bursts of rapid, poorly articulated speech (20). The phenotype varies with the occurrence of co-morbid conditions, with lower IQ and more severe autism symptoms associated with more serious language problems (16,21–25).

The working group considered a number of measures, including standardized tests, caregiver reports, and laboratory-based tasks. Significant limitations of most measures were noted, including a lack of sensitivity to clinically meaningful change and limited or unknown reliability and validity (26). In the end, the working group concluded that expressive language sampling procedures held the most promise for immediate use in clinical trials (see Table 2). These procedures are common in research settings and involve recording samples of spontaneously produced language in a structured social interaction with an examiner (26). Objective variables can be created from transcripts of these samples to reflect numerous aspects of the linguistic phenotype. These procedures are applicable (with minor modifications) over a wide developmental range (26).

There is considerable evidence that measures derived from language samples collected under standardized conditions are related to chronological age in typically developing children and mental age in many language-disordered populations. These measures distinguish language-typical from language-disordered populations (27) and evidence is accumulating that these measures can distinguish individuals with FXS from those with other neurodevelopmental disorders, such as Down syndrome (13,22,28–31). Guidelines for standardization for a number of language sampling procedures (e.g., conversation, narration) exist (32), and these procedures can be mastered with rather minimal training (26).

Memory and Learning

The working group distinguished between several types of memory based on the duration and type of information stored, as well as the neural systems involved (33–39). In terms of *explicit memory*, which refers to those aspects of memory that involve effortful attempts to store and retrieve information, individuals with FXS are more impaired in remembering sequentially presented information than simultaneously presented information (40). Individuals with FXS also perform below expectations on measures of *working memory* (41), which involves rehearsal and allocation of mental resources (42,43). Individuals with FXS also are especially impaired at working memory tasks requiring the manipulation of auditory (as opposed to visual) information (44–46) and when memorization requires considerable attention and strategic planning (47,48). The *implicit memory* system, which involves the storage of information without conscious effort, is largely unexplored in FXS, although anecdotal experience suggests that implicit memory may not be as impaired as other types of memory in FXS.

The working group identified several memory tasks that mapped onto the FXS memory phenotype (see Table 2), for example, the Auditory Working Memory and Digits Reversed subtests of the Woodcock-Johnson III Tests of Cognitive Abilities (49,50), although the latter subtest may be beyond the ability of individuals with a developmental level less than four years. The Corsi Blocks (51) test has been used successfully to study visual-spatial sequential memory in FXS. The working group also proposed assessing long-term memory with the list learning and story memory subtests of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; 52, 53).

Executive Functioning—Executive functioning involves setting goals, developing plans, and implementing goals (54). The functioning of this system is typically indexed by tasks that require behaviors such as inhibition of prepotent responses, monitoring the success of a problem-solving strategy, the ability to shift between responses or strategies, and the efficiency of applying cognitive processes (55). There is compelling evidence that deficits in executive functioning are highly characteristic of individuals with FXS (56,57), with performance often below mental age expectations (58). These impairments in executive function map on to abnormalities in brain function, particularly in frontal regions in FXS (59,60).

The working group identified several measures of executive function that appear well-suited for use in some clinical trials targeting FXS (see Table 2), however, some of these measures may be beyond the ability of lower-functioning individuals. The working group also suggested the use of informant report measures (e.g., the Behavior Rating Inventory of Executive Function; BRIEF; 61), which may prove useful when direct assessment is not possible.

Social Cognition—Social cognition refers to the processing of information about the social world (e.g., the ability to recognize differences in people’s knowledge or perspectives). These skills are impaired relative to chronological age expectations in FXS (62–64). Impairments in some aspects of social cognition exceed the impairments seen in other domains, especially in individuals with co-morbid FXS and autism (23).

The working group noted that measures of social cognition have not been well-studied for FXS. More generally, it was noted that the relationship between performance on such measures and actual social behavior is not strong. The working group concluded that the Social Responsiveness Scale (65), which is a caregiver report, and the use of eye-tracking technology to assess attention to social events, were promising but required further validation in FXS.

Academic Achievement—The working group recommended the use of academic achievement measures as indicators of long-term outcomes and real-life functioning, particularly in clinical trials of longer duration. The expectation is that changes in language, memory, and executive function are likely to lead to functional changes in the academic domain. However, no specific measure was found to be adequate at this time.

Summary—Although promising measures were identified for several key domains, each requires further evaluation before being endorsed for clinical trials.

- Additional data are needed on test-retest reliability, sensitivity, and validity.
- Validation over an expanded age range is needed.
- Feasibility and validity data are needed across the full range of affectedness.

Behavior and Emotion Measures

The working group on Behavior and Emotion considered the following points in their evaluation of measures that could be appropriate for use in a clinical trial: 1. The measure should adequately cover the type and severity of problems that most interfere with daily functioning and quality of life for individuals with FXS and their families. In this regard, the FDA clearly emphasizes that a patient reported outcome (PRO) measure, such as a rating scale, must include symptoms that affected individuals or their proxies (caregivers), clinicians, and investigators report to be characteristic of the specific condition to be treated (66). 2. The measure should have documented reliability and validity for the age range and level of functioning of individuals likely to be recruited into clinical trials, and preferably the psychometric properties should be established for the FXS population rather than for intellectual disability more generally. 3. The measure should have a proven capacity to detect improvement in the specified domain (i.e., sensitivity), preferably in the context of a controlled trial. The working group identified several maladaptive behaviors that are over-represented in FXS relative to other intellectual disabilities: 1) inattention; 2) hyperactivity/impulsivity; 3) irritability/aggression; 4) self-injury; 5) anxiety; 6) repetitive/compulsive behavior; 7) sleep problems; and 8) social avoidance/reciprocity. Adaptive behavior (the ability to complete the activities necessary for successful navigation of independent daily life) was also identified as a potential target. A list of promising measures of maladaptive behaviors can be found in Table 3.

The working group considered informant-based rating scales, such as those completed by a parent, caregiver, or clinician, because of their ease of use, cost, and utility in multi-site trials. Some of these measures have a track record documenting behavioral improvement in individuals with intellectual disability or autism spectrum disorders, such as the Aberrant Behavior Checklist (67–69). The working group also acknowledged several limitations and

concerns about informant reports, including a reliance on respondents who may be unintentionally biased such as parents of affected children. In addition, informant-based rating scales are quite variable with some rating *severity of symptoms* and others rating *frequency*. In addition, well-defined “behavioral anchors” to help guide respondents are often absent. This can be problematic for caregivers, many of whom have minimal exposure to other children with FXS and thus, may have little experience with the full range of symptom severity. The working group concluded, however, that the sensitivity of these rating scales to treatment effects is likely to be dependent on test-retest reliability within-respondent and within-subject. Thus, such rating scales remain a viable option, especially given the paucity of other validated measures.

A challenge for any rating scale is providing coverage of the wide range of symptom severity and expression across individuals with FXS. An additional challenge arises in covering the variability in symptom presentation across developmental and mental ages.

The working group identified several behavioral rating scales that meet important criteria, among these is the Aberrant Behavior Checklist (67), which has good psychometric properties when used with individuals with intellectual disability and has been used successfully to document improvements in irritability/aggression in many controlled trials of autism (70–73). Berry-Kravis et al (72) and Fung et al. (73) reported on the successful use of the ABC total score to document behavioral improvement in an open label trial of lithium in 15 children and adults with FXS and donepezil in 12 children and adults with FXS, respectively.

At the time of the working group meeting, the ABC had not been extensively validated for FXS and thus the working group endorsed further study on this measure, especially in light of the fact that several ongoing industry-sponsored clinical trials have utilized the ABC as a primary endpoint for FXS. This recommendation led to a successful five-site collaborative study of the psychometric properties of the ABC in over 600 individuals with FXS (74) and utilization in clinical trials (75,76). The working group also identified both the SNAP-IV (77) and the Connors Rating Scales (78), each designed to measure hyperactivity/impulsivity, as having utility in clinical trials for FXS, however it was noted that only limited data are available for those with intellectual disabilities. Although anxiety is an important problem in FXS, the working group concluded that few suitable measures are currently available. The Anxiety, Depression and Mood Screen (ADAMS) (79) was identified as a useful tool for screening anxiety and mood symptoms in individuals with intellectual disabilities, and has since shown evidence of validity for FXS (80); however, its sensitivity to treatment effects is unknown. The working group also saw promise in the Pediatric Anxiety Rating Scale (PARS) (81), which was used successfully to document improvement in pediatric anxiety disorders with SSRIs. A small reliability and validity study of the PARS for FXS has been completed recently and is in press (82).

Repetitive behaviors, such as perseverative speech and stereotypies, are often problematic in FXS. The Children’s Yale Brown Obsessive Compulsive Scale, modified for pervasive developmental disorders (CYBOCS-PDD) (83), has excellent psychometric properties when used with children with an autism spectrum disorder and has documented sensitivity to psychopharmacological intervention (84), although its reliability, validity and sensitivity in FXS is unknown. The Repetitive Behavior Scale-Revised (RBS-R) (85,86) was designed for use in populations with an autism spectrum disorder and related developmental disorders and was recently used in a study of young boys with FXS or idiopathic autism (87), with different profiles of repetitive behaviors being observed in the two groups.

Although the main focus in clinical trials in FXS has been on maladaptive behaviors shown to improve with short-term treatment in other disorders, it is recognized that adaptive behavior may also be responsive to intervention, especially in studies providing treatment over extended periods. Adaptive behavior scales are expected to be especially important in assessing treatment impact in very young children with FXS, as fewer measures of cognitive function exist. It is also possible that older children and adults may demonstrate improvements in adaptive behaviors (e.g., improved communication, greater independence with daily living skills, or improved socialization) that are secondary to reductions in maladaptive behaviors or enhanced cognition. The relative value of different adaptive behavior measures has not been explored extensively in FXS. However, improvement in adaptive behavior using specific subscales of the Vineland (88) has been demonstrated in the small open label trial of targeted treatment with lithium (Expressive Language subdomain) (72), as well as the recent placebo-controlled trial of arbaclofen (Play and Leisure subdomain) (75), suggesting that the measure is likely to be sensitive.

In light of the limitations of informant-report rating scales, the working group also concluded that there is a pressing need for the development of more objective, direct observation or assessment of individuals with FXS by trained clinicians or experimenters. Such measures could include behavioral sampling, in which individuals with FXS are observed directly by trained and reliable experimenters across several contexts. The advantages of direct behavioral observation should be balanced, however, by the time commitment and potential cost associated with such an approach for a large multi-site trial.

The working group also discussed the feasibility and advantages of developing a new behavioral scale specifically designed for individuals with FXS. It was suggested that a measure could be developed by combining items selected from other instruments or developed anew by a diverse panel of parents, clinicians, and investigators, and empirically developed using established methods of factor analytic techniques, as well as reliability and validity studies. Such an approach could yield a tool measuring behaviors closer to the “core” of the FXS phenotype and a scale or scales that would be more sensitive to interventions designed to target the neurobiology of the disorder.

Summary—The working group endorses efforts:

- To establish the reliability, validity and sensitivity (including sensitivity to change) of several currently available measures within the FXS population
- To establish the specificity of the measured constructs to FXS by examination of correlations between such measures and valid FXS biomarkers (e.g., brain imaging) and gene-dose (e.g. FMRP)
- To consider development of a new behavior rating scale for FXS to cover the phenotype and the full range of associated symptoms
- To supplement traditional psychometric studies with data from focus groups that include patients or their proxies, and other caregivers to provide input on the construct validity, interpretability, and feasibility of measures of interest

Biomarkers and Medical Measures

Biomarkers as endpoints for treatment response are attractive because of their potential to be objective, repeatable, and quantifiable measures of the biology of the disorder or a physical manifestation of the underlying biology. Such biomarkers might assess the impact of a treatment on mechanisms underlying FXS as opposed to evaluating behavioral symptoms alone. In practice, however, it has been difficult to ascertain whether and how changes in

available biomarkers correlate with meaningful improvements in function or behavior (89). Thus, the working group concluded that our understanding of biomarkers in FXS is not sufficiently developed to be useful for product registration with the FDA at this time. The need for clinically meaningful, treatment-sensitive biomarkers in FXS, however, cannot be over-stated.

Blood and Tissue Biomarkers—Research on peripheral tissue biomarkers in FXS was noted by the working group to be at a very early stage. *FMR1* methylation status, *FMR1* mRNA levels and FMRP levels have been proposed as obvious markers of the disorder but may or may not predict response to disease-targeted pharmacological interventions, as they do not represent the target of current therapies. In a small pilot trial of mGluR5 blocker AFQ056, response was substantially better in the sub-group of FXS males with complete gene methylation and undetectable *FMR1* mRNA levels (76)(see Table 1). These results need to be replicated in a larger trial, but suggest that *FMR1* methylation and FMRP (not measured in the AFQ056 study) should continue to be studied as biomarkers that may help predict treatment response in future FXS trials.

Although predictors of treatment response are critical, a major focus of this working group was on development of measures that reflect the impact of a drug on FMRP-regulated pathways, including activity of 1) proteins in the translational activation pathway and 2) proteins whose translation is regulated by FMRP (some proteins may fit in both of these categories). ERK activation rate, a translational pathway biomarker, abnormal in FXS mice and humans, is responsive (normalizes) to lithium (72) and riluzole (90) treatment in individuals with FXS (see Table 1). Amyloid precursor protein (APP), a potential biomarker whose translation is regulated by FMRP (91), is a target measure in several FXS trials; for example, APP was shown to normalize during treatment with acamprosate (Table 1). Other measures, such as phosphoinositide 3-kinase (92), mTOR and S6 kinase activity (93), are promising, but need further development and validation for meaningful implementation.

In the development of peripheral biomarkers, it is important to recognize that pathways and receptors are not always connected in the same fashion in different tissues. For biomarker discovery, careful consideration must be given to: 1) tissue differences in the interaction of the biomarker with pathways affected by FMRP and the drug under study; 2) individualization of biomarkers for agents with different targets; 3) peripheral biomarkers index tissues outside the CNS and may be very difficult to correlate with a CNS-mediated behavioral or functional improvement. Consequently, biomarkers in FXS may be, at best, an indication that the drug is hitting a desired cellular target, but may not fully predict treatment response. In addition, a reasonable goal could be to identify surrogate biomarkers that reliably track with the disorder or treatment regimen, but may not be central to the biological pathways in FXS.

Electrophysiological Measures—Prepulse inhibition (PPI) of the startle response, the most extensively studied electrophysiological measure, is reduced in individuals with FXS and has good test-retest reproducibility (94). Moreover, PPI is attractive because (a) mGluR5 pathways are involved in both PPI regulation and FXS, (b) there are prominent sensory inhibitory (gating) deficits in FXS, and (c) PPI appears deficient in both the *Fmr1* knockout mouse and in humans with FXS (95,96) (see Table 1). There are problems, however, with establishing functional behavioral correlates for PPI, and achieving inter-site reliability. Consequently, PPI may be most useful in early proof-of-concept trials as opposed to large multi-site trials.

Autonomic measures including heart rate, R-R interval, and beat-to-beat variability were not found to be responsive to lithium treatment but have not been utilized in other treatment

studies (97). Other psychophysiological measures are under investigation as potential biomarkers for treatment response in individuals with FXS.

Eye Tracking and Pupillometry—Eye gaze aversion and pupillometry measured with newer eye tracking systems represent potential biomarkers of anxiety and excess sympathetic outflow. Both eye gaze and pupillometry measures are abnormal in FXS (98) and have good test-retest reproducibility (99). This measure is being used in clinical trials currently (Table 1) but data are not yet available on sensitivity to treatment-induced change.

Although the behavioral correlates of eye gaze aversion and social anxiety would seem obvious, it has been difficult to show correlations of pupillometry and eye tracking response profiles with validated instruments for rating such behaviors and thus, additional research is needed before eye tracking could be considered a possible endpoint supporting registration of pharmaceuticals for treatment of FXS. Moreover, the cost of eye tracking systems may limit use of these measures to smaller proof-of-concept trials.

Neuroimaging Studies—Thus far, early phase clinical trials in FXS have utilized behavior checklists or cognitive tests developed for general use in non-FXS populations, as clinical trial endpoints. Because neurobiological mechanisms associated with FXS interact with environmental influences to affect brain development and function from birth, the use of relatively non-specific behavioral or cognitive endpoints increases the likelihood of a false negative (Type II) error, even in the presence of a biologically meaningful effect.

Because direct measures of brain structure, function and neurochemistry might be more closely and directly related to fundamental neurobiological etiology arising from reduced FMRP, the possibility of employing in vivo neuroimaging biomarkers as disease biomarkers or treatment endpoints in FXS has attracted considerable interest (100). A growing number of brain imaging studies indicate that distinct and, in some cases, unique neuro-phenotypic features exist in individuals with FXS relative to age and sex matched neuro-typical controls, as well as IQ matched idiopathic (i.e., non-FXS) developmentally disabled or autistic controls (101–108). These features include a significantly (40%) larger caudate nucleus (a prominent finding detectable as early as one year of age), aberrant white matter connectivity or activation of prefrontal-striatal pathways, and significant anatomical and/or functional activation differences in the insular, fusiform and superior temporal cortices, amygdala, hippocampus and cerebellar vermis.

The few longitudinal neuroimaging studies conducted in children and adolescents to date indicate that the trajectory of brain development in FXS diverges from typical neurodevelopment at distinct time periods during childhood and adolescence as well as in young adults (109,110). These critical windows of brain maturation may correspond to neurodevelopmental epochs when FMRP is particularly important for achieving optimal synaptic maturation and connectivity in the brain and could in the future, be used to evaluate treatment outcome.

An initial magnetic resonance spectroscopy (MRS) study in a small sample of individuals with FXS showed reduced choline levels in a region of the prefrontal cortex (111). A subsequent study replicated this finding in the caudate nucleus of individuals with FXS relative to IQ-matched controls and also found reduced levels of GLX, an MRS measure comprised of glutamine and glutamate (submitted). Because MRS techniques can detect and measure several neurometabolites implicated in the neurobiology of FXS, this neuroimaging modality holds particular promise for identifying potential disease biomarkers and clinical trial endpoints.

In summary, because of the seemingly greater proximity to fundamental mechanisms underlying FXS, neuroimaging biomarkers may hold promise as potential endpoints in clinical trials. Notwithstanding the cost and training involved, additional research is needed to demonstrate the feasibility and utility of using neuroimaging biomarkers in research and clinical settings for this purpose.

Side Effect Monitoring, Side effects may be difficult to detect in cognitively impaired individuals who may not be able to report symptoms accurately. Formal comprehensive questionnaires have been used such as the SMURF (72); however, these are time consuming and some studies have used lists for screening of an inventory of specific targeted side effects (100) depending on the particular concerns associated with the drug being tested. Considerable care should be exercised in this arena to monitor adverse events and protect study subjects.

Summary

- Research on biomarkers for detecting intervention-induced improvement in fundamental neurobiological deficits of FXS is in its infancy.
- Continued research is greatly needed at multiple levels: cellular, physiological, neuroimaging and (core) behavioral.
- Measures under development have yet to be directly linked to the underlying neurobiology of FXS.
- The future availability of robust biomarkers for elucidating the neurobiology of, and treatment-response in humans with FXS is of utmost importance.

Discussion

The overarching objective of these NIH-initiated meetings was to facilitate discussion that would ultimately lead to treatment outcome measures that can be utilized in current and future clinical trials. Particularly motivating to the participants was the knowledge that a core set of relevant clinical endpoints would significantly improve our ability to interpret and compare results from clinical trials taking place at different sites or using different pharmacologic agents and/or behavioral and cognitive interventions. This may also increase the likelihood that future Phase 3 trials would result in more rapid FDA approval of drugs for indications related to FXS with labeling of potential therapeutic agents for clinical use. The process may also serve as a model for other neurodevelopmental disorders.

The working group suggests that no single measure currently exists that meets all criteria for an ideal clinical endpoint that can be used to evaluate treatment for FXS. Although the ABC is proving quite useful in current clinical trials, it is unlikely to meet the long-term needs of the field given the limitations noted in the Results section. The working groups endorsed the need to devise a set of measures that reliably captures the core cognitive impairments and underlying neurobiological mechanisms of FXS, as well as the impairments that define key behavioral and emotional domains. Consequently, we conclude that:

1. Scientists, funding agencies, and private industry should collaborate to support a broad-based program of research in FXS to arrive at a set of core measures tapping the domains of cognition, behavior and emotion, and neurobiology.
2. This research should strive to identify and develop measures that (a) are psychometrically sound with high reliability, validity, and sensitivity to change for the FXS population; (b) index core features of the FXS phenotype, at the neurobiological level or the level of cognition, emotion, and behavior; (c) focus on

constructs that translate to improved quality of life for individuals with FXS; (d) are feasible across a wide range of ages and levels of functional impairment or have well-documented conditions of use with respect to age and degree of impairment; (e) are relatively insensitive to placebo effects and (f) are logistically feasible for use in large multisite trials.

3. Several measures appear particularly promising and thus, should be a high priority for further development. Arguably, the single measure that appears closest to reaching the threshold of the criteria identified is the Aberrant Behavior Checklist, which has a long history of use in research, including clinical trials and has now been revised to better accommodate the FXS phenotype. Importantly, the ABC is being used in clinical trials targeting FXS and is proving useful to detecting treatment-induced change. In the cognitive domain, expressive language sampling holds promise and has been refined in ways that make it feasible for use in multisite clinical trials with individuals affected by FXS. In the domain of biomarkers, PPI and neuroimaging appear promising at least for proof-of-concept trials and perhaps more broadly.
4. Rather than begin anew, there are measures that have been (or are being) developed for other clinical populations and many of the constructs being measured overlap the core symptoms of FXS. These measures should be evaluated for their feasibility and face validity for FXS. This approach of “borrow, adapt, and evaluate” should prove useful for all domains, including that of biomarkers
5. The development of outcome measures would benefit from greater collaboration between scientists working on different model systems. There is currently a disconnect between efforts to measure the phenotypes of FXS in humans and those in mice and other non-human model systems. An attempt to create parallel measures that index similar constructs across different model systems is likely to assist in measure development and accelerate translation of therapeutics from preclinical work to human clinical trials, and has been successful in other disorders.
6. Although FDA approval of a drug is likely to require selection of a single measure in advance of the trial, it is important to acknowledge that our understanding of the causal pathways from genetic mutation to individual phenotype in FXS is incomplete. Thus, a useful strategy would be to select a small set of core measures, with the goal being to advance understanding of FXS behavioral domains and biomarkers and the targeted neurobiological systems and pathways.. In this way, we will document efficacy of the therapeutic agent and simultaneously provide new insights into the pathophysiology of FXS. Ultimately, however, we must arrive at a primary measure for any given trial, although this measure is likely to be different for different therapeutic agents or for short- versus long-term trials, with the choice being informed by knowledge about the causal pathways of the disorder and the mechanisms of action attributed to the therapeutic agent.

In conclusion, this is an exciting era in the field of FXS. The interest in potential therapeutics has increased dramatically and exciting partnerships are being forged between academia, government agencies, and private industry with the goal of moving us closer to a cure. Moreover, several promising agents are already available for testing, with some already have documented positive effects. As a result, optimism among families affected by FXS is high. Further progress will be limited, however, unless the field can quickly arrive at a set of outcome measures to be used in clinical trials. If we do not, we run the risk of abandoning potentially effective treatments and adding to the stress and sense of loss of affected families. The stakes could not be higher.

Acknowledgments

The Fragile X Outcome Measures Working Group meeting was supported by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), the Office of Rare Disease Research (ORDR), the National Institute of Mental Health (NIMH), and the National Institute of Neurological Disorders and Stroke (NINDS). The authors thank Dr. Shaguna Mathur for her assistance with manuscript formatting and submission.

References

1. Hagerman PJ. The fragile X prevalence paradox. *J Med Genet.* 2008; 45:498–499. [PubMed: 18413371]
2. Fernandez-Carvajal I, Walichiewicz P, Xiaosen X, et al. Screening for expanded alleles of the FMR1 gene in blood spots from newborn males in a Spanish population. *J Mol Diagn.* 2009; 11:324–329. [PubMed: 19460941]
3. Grossman AW, Aldridge GM, Weiler IJ, et al. Local protein synthesis and spine morphogenesis: fragile X syndrome and beyond. *J Neurosci.* 2006; 26:7151–7155. [PubMed: 16822971]
4. Bassell GJ, Warren ST. Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function. *Neuron.* 2008; 60:201–214. [PubMed: 18957214]
5. Darnell JC, Van Driesche SJ, Zhang C, et al. FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell.* 2011; 146:247–261. [PubMed: 21784246]
6. Bhakar AL, Dölen G, Bear MF. The pathophysiology of fragile X (and what It teaches us about synapses). *Annu Rev Neurosci.* 2012; 35:417–443. [PubMed: 22483044]
7. Hagerman RJ, Berry-Kravis E, Ono MY, et al. Advances in the treatment of fragile X syndrome. *Pediatrics.* 2009; 123:378–390. [PubMed: 19117905]
8. Berry-Kravis E, Potanos K. Psychopharmacology in fragile X syndrome--present and future. *Ment Retard Dev Disabil Res Rev.* 2004; 10:42–48. [PubMed: 14994287]
9. Berry-Kravis E, Sumis A, Hervey C, et al. Clinic-based retrospective analysis of psychopharmacology for behavior in fragile X syndrome. *Int J Peds.* 2012 epub July 30.
10. Berry-Kravis E, Knox A, Hervey C. Targeted treatments for fragile X syndrome. *J Neurodev Disord.* 2011; 3:193–210. [PubMed: 21484200]
11. Hall SS, Hammond JL, Hirt M, et al. A 'learning platform' approach to outcome measurement in fragile X syndrome: a preliminary psychometric study. *J Intellect Disabil Res.* 2012 epub April 25.
12. Abbeduto L, Brady N, Kover ST. Language development and fragile X syndrome: Profiles, syndrome-specificity, and within-syndrome differences. *Ment Retard Dev Disabil Res Rev.* 2007; 13:36–46. [PubMed: 17326110]
13. Price JR, Roberts JE, Hennon EA, et al. Syntactic complexity during conversation of boys with fragile X syndrome and Down syndrome. *J Speech Lang Hear Res.* 2008; 51:3–15. [PubMed: 18230852]
14. Abbeduto L, Murphy MM, Richmond EK, et al. Collaboration in referential communication: Comparison of youth with Down syndrome or fragile X syndrome. *Am J Ment Retard.* 2006; 111:170–183. [PubMed: 16597184]
15. Abbeduto L, Murphy MM, Kover ST, et al. Signaling noncomprehension of language: A comparison of fragile X syndrome and Down syndrome. *Am J Ment Retard.* 2008; 113:214–230. [PubMed: 18407723]
16. Estigarribia B, Martin GE, Roberts JE, et al. Narrative skill in boys with fragile X syndrome with and without autism spectrum disorder. *Appl Psycholinguist.* 2011; 32:359–388. [PubMed: 21516264]
17. Roberts J, Martin GE, Moskowitz L, et al. Discourse skills of boys With fragile X syndrome in comparison to boys with Down syndrome. *J Speech Lang Hear Res.* 2007; 50:475–492. [PubMed: 17463242]
18. Belser RC, Sudhalter V. Conversational characteristics of children with fragile X syndrome: repetitive speech. *Am J Ment Retard.* 2001; 106:28–38. [PubMed: 11246710]

19. Sudhalter V, Cohen IL, Silverman W, et al. Conversational analyses of males with fragile X, Down syndrome, and autism: Comparison of the emergence of deviant language. *Am J Ment Retard.* 1990; 94:431–44. [PubMed: 2137003]
20. Roberts JE, Hennon EA, Price JR, et al. Expressive language during conversational speech in boys with Fragile X syndrome. *Am J Ment Retard.* 2007; 112:1–17. [PubMed: 17181388]
21. Budimirovic DB, Bukelis I, Cox C, et al. Autism spectrum disorder in Fragile X syndrome: differential contribution of adaptive socialization and social withdrawal. *Am J Med Genet A.* 2006; 140:1814–1826. [PubMed: 16906564]
22. Kover ST, Abbeduto L. Expressive language in male adolescents with fragile X syndrome with and without comorbid autism. *J Intellect Disabil Res.* 2010; 54:246–265. [PubMed: 20146742]
23. Lewis P, Abbeduto L, Murphy MM, et al. Cognitive, language, and social-cognitive skills of individuals with fragile X syndrome with and without autism. *J Intellect Disabil Res.* 2006; 50:532–545. [PubMed: 16774638]
24. McDuffie A, Kover ST, Abbeduto L, et al. Profiles of receptive and expressive language in males with comorbid fragile X syndrome and autism. *Am J Intellect Dev Disabil.* 2012; 117:18–32. [PubMed: 22264110]
25. Philofsky A, Hepburn SL, Hayes A, et al. Linguistic and cognitive functioning and autism symptoms in young children with Fragile X syndrome. *Am J Ment Retard.* 2004; 109:208–218. [PubMed: 15072521]
26. Abbeduto, L.; Kover, ST.; McDuffie, A. Studying the language development of children with intellectual disabilities. In: Hoff, E., editor. *Handbook of child language research methods.* West Sussex, UK: Wiley-Blackwell; 2012. p. 330–346.
27. Heilmann JJ, Miller JF, Nockerts A. Using language sample databases. *Language, Speech, and Hearing Services in Schools.* 2010; 41:84–95.
28. Estigarribia B, Martin GE, Roberts JE. Cognitive, environmental, and linguistic predictors of syntax in fragile X syndrome and Down syndrome. *J Speech Lang Hear Res.* 2012 epub Apr 3.
29. Finestack LH, Abbeduto L. Expressive language profiles of verbally expressive adolescents and young adults with Down syndrome or fragile X syndrome. *J Speech Lang Hear Res.* 2012; 53:1334–1348. [PubMed: 20643789]
30. Finestack LH, Palmer M, Abbeduto L. Macrostructural narrative language of adolescents and young adults with Down syndrome or fragile X syndrome. *Am J Speech Lang Pathol.* 2012; 21:29–46. [PubMed: 22049405]
31. Finestack LH, Sterling A, Abbeduto L. Discriminating Down syndrome and fragile X syndrome based on language ability. *Journal Child Lang.* 2012 in press.
32. Abbeduto L, Benson G, Short K, et al. Effects of sampling context on the expressive language of children and adolescents with mental retardation. *Ment Retard.* 1995; 33:279–288. [PubMed: 7476250]
33. Reber, AS. *Implicit learning and tacit knowledge: An essay on the cognitive unconscious.* London: Oxford University Press; 1993.
34. Siegert RJ, Taylor KD, Weatherall M, et al. Is implicit sequence learning impaired in Parkinson's disease? A meta-analysis. *Neuropsychology.* 2006; 20:490–495. [PubMed: 16846267]
35. Baddeley, AD. *Working memory.* Oxford: Oxford University Press; 1986.
36. Baddeley AD. Working memory and language: An overview. *J Commun Disord.* 2003; 36:189–208. [PubMed: 12742667]
37. Jarrold C, Baddeley AD. Short-term memory for verbal and visuo-spatial information in Down's syndrome. *Cogn Neuropsychiatry.* 1997; 2:101–122.
38. D'Esposito M, Aguirre GK, Zarahn E, et al. Functional MRI studies of spatial and nonspatial working memory. *Cog Brain Res.* 1998; 7:1–13.
39. Walter H, Bretschneider V, Gron G, et al. Evidence for quantitative domain dominance for verbal and spatial working memory in frontal and parietal cortex. *Cortex.* 2003; 39:897–911. [PubMed: 14584558]
40. Dykens EM, Hodapp RM, Leckman JF. Strengths and weaknesses in the intellectual functioning of males with fragile X syndrome. *Am J Ment Defic.* 1987; 92:234–236. [PubMed: 3434594]

41. Ornstein PA, Schaaf JM, Hooper SR, et al. Memory skills of boys with fragile X syndrome. *Am J Ment Retard.* 2008; 113:453–465. [PubMed: 19127656]
42. Acheson DJ, MacDonald MC. Verbal working memory and language production: common approaches to the serial ordering of verbal information. *Psychol Bull.* 2009; 135:50–68. [PubMed: 19210053]
43. Cowan, N. *Attention and memory: an integrated framework.* Oxford, England: Oxford University Press; 1995.
44. Baker S, Hooper S, Skinner M, et al. Working memory subsystems and task complexity in young boys with fragile X syndrome. *J Intellect Disabil Res.* 2011; 55:19–29. [PubMed: 21121991]
45. Cornish KM, Kogan CS, Li L, et al. Lifespan changes in working memory in fragile X premutation males. *Brain Cogn.* 2009; 69:551–558. [PubMed: 19114290]
46. Munir F, Cornish KM, Wilding J. A neuropsychological profile of attention deficits in young males with fragile X syndrome. *Neuropsychologia.* 2000; 38:1261–1270. [PubMed: 10865102]
47. Cornish KF, Munir F, Wilding J. A neuropsychological and behavioural profile of attention deficits in fragile X syndrome. *Rev Neurol.* 2001; 33:S24–29. [PubMed: 12447815]
48. Lanfranchi SO, Jerman E, Dal Pont A, et al. Executive function in adolescents with Down Syndrome. *J Intellect Disabil Res.* 2012; 54:308–319. [PubMed: 20202074]
49. Woodcock, Richard W.; Mather, Nancy; McGrew, Kevin S. *Woodcock–Johnson III Tests of Cognitive Abilities Examiner’s Manual.* Itasca, IL: Riverside; 2001.
50. Woodcock, Richard R.; McGrew, Kevin S.; Schenk, Fredrick A. *Woodcock–Johnson III Normative Update Technical Manual.* Itasca, IL: Riverside; 2007.
51. Kessels RPC, van Zandvoort MJE, Postma A, et al. The Corsi Block-Tapping Task: Standardization and Normative Data. *Applied Neuropsychology.* 2000; 7:252–258. [PubMed: 11296689]
52. Randolph, C. *Repeatable Battery for the Assessment of Neuropsychological Status.* San Antonio, TX: Pearson; 1998.
53. Berry-Kravis E, Lara R, Kim O-K, et al. Characterization of potential outcome measures for future clinical trials in fragile X syndrome. *J Autism Dev Disord.* 2008; 38:1751–1757. [PubMed: 18369716]
54. Welsh MC, Pennington BF. Assessing frontal lobe functioning in children: Views from developmental psychology. *Dev Neuropsychol.* 1988; 4:199–230.
55. Pennington BF. Using genetics to dissect cognition. *Am J Hum Genet.* 1997; 60:13–16. [PubMed: 8981941]
56. Wilding JK, Cornish K, Munir F. Further delineation of the executive deficit in males with fragile-X syndrome. *Neuropsychologia.* 2002; 40:1343–1349. [PubMed: 11931937]
57. Scerif GK, Cornish K, Wilding J, et al. Delineation of early attentional control difficulties in fragile X syndrome: focus on neurocomputational changes. *Neuropsychologia.* 2007; 45:1889–1898. [PubMed: 17254617]
58. Hooper SR, Hatton D, Sideris J, et al. Executive functions in young males fragile X syndrome in comparison to mental age-matched with controls: baseline findings from a longitudinal study. *Neuropsychology.* 2008; 22:36–47. [PubMed: 18211154]
59. Kwon H, Menon V, Eliez S, et al. Functional neuroanatomy of visuospatial working memory in fragile X syndrome: relation to behavioral and molecular measures. *Am J Psychiatry.* 2001; 158:1040–1051. [PubMed: 11431225]
60. Tamm L, Menon V, Johnston CK, et al. fMRI study of cognitive interference processing in females with fragile X syndrome. *J Cogn Neurosci.* 2002; 14:160–171. [PubMed: 11970783]
61. Gioia G, Isquith PK, Guy SC, et al. Behavior Rating Inventory of Executive Function. *Child Neuropsychology.* 2000; 6:235–238. [PubMed: 11419452]
62. Cornish K, Burack JA, Rahman A, et al. Theory of mind deficits in children with fragile X syndrome. *J Intellect Disabil Res.* 2005; 49:372–8. [PubMed: 15817054]
63. Garner C, Callias M, Turk J. Executive function and theory of mind performance of boys with fragile-X syndrome. *J Intellect Disabil Res.* 1999; 43:466–74. [PubMed: 10622362]

64. Mazzocco MM, Pennington BF, Hagerman RJ. Social cognition skills among females with fragile X. *J Autism Dev Disord.* 1994; 24:473–85. [PubMed: 7961331]
65. Constantino, JN.; Gruber, CP. *The Social Responsiveness Scale Manual.* California: Western Psychological Services; 2005.
66. U.S. Department of Health and Human Services. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2006.
67. Aman M, Singh N, Stewart A, et al. The Aberrant Behavior Checklist: A behavior rating scale for the assessment of treatment effects. *Am J Ment Defic.* 1985; 89:485–491. [PubMed: 3993694]
68. McCracken JT, McGough J, Shah B, et al. Risperidone in children with autism and serious behavioral problems. *N Engl J Med.* 2002; 347:314–321. [PubMed: 12151468]
69. Zarcone JR, Hellings JA, Crandall K, et al. Effects of risperidone on aberrant behavior of persons with developmental disabilities: I. A double-blind crossover study using multiple measures. *Am J Ment Retard.* 2001; 106:525–38. [PubMed: 11708938]
70. Arnold LE, Vitiello B, McDougle C, et al. Parent-defined target symptoms respond to risperidone in RUPP autism study: customer approach to clinical trials. *J Am Acad Child Adolesc Psychiatry.* 2003; 42:1443–50. [PubMed: 14627879]
71. McDougle CJ, Scahill L, McCracken JT, et al. Research Units on Pediatric Psychopharmacology (RUPP) Autism Network. Background and rationale for an initial controlled study of risperidone. *Child Adolesc Psychiatr Clin N Am.* 2000; 9:201–224. [PubMed: 10674197]
72. Berry-Kravis E, Sumis A, Hervey C, et al. Open-label treatment trial of lithium to target the underlying defect in fragile X syndrome. *J Dev Behav Pediatr.* 2008; 29:293–302. [PubMed: 18698192]
73. Fung LK, Quintin E, Haas BW, et al. Conceptualizing neurodevelopmental disorders through a mechanistic understanding of fragile X syndrome and Williams syndrome. *Curr Opin Neurol.* 2012; 25:112–124. [PubMed: 22395002]
74. 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; 42:1377–1392. [PubMed: 21972117]
75. Berry-Kravis E, 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; 4:152ra127.
76. Jacquemont S, Curie A, des Portes V, et al. Epigenetic modification of the FMR1 gene in fragile X syndrome is associated with differential response to the mGluR5 antagonist AFQ056. *Sci Transl Med.* 2012; 3:64ra1.
77. Swanson JM, Kraemer HC, Hinshaw SP, et al. Clinical relevance of the primary findings of the MTA: success rates based on severity of ADHD and ODD symptoms at the end of treatment. *J Am Acad Child Adolesc Psychiatry.* 2001; 40:168–79. [PubMed: 11211365]
78. *Conners' Rating Scales-Revised Technical Manual.* North Tonawanda, New York: Multi Health Systems; 2000.
79. Esbensen AJ, Rojahn J, Aman MG, et al. Reliability and validity of an assessment instrument for anxiety, depression, and mood among individuals with mental retardation. *J Autism Dev Disord.* 2003; 33:617–629. [PubMed: 14714931]
80. Cordeiro L, Ballinger E, Hagerman RJ, et al. Clinical assessment of DSM-IV anxiety disorders in fragile X syndrome: Prevalence and characterization. *J Neurodevelopmental Disord.* 2011; 3:57–67.
81. Riddle MA, Ginsburg GS, Walkup JT, et al. The pediatric anxiety rating scale (PARS): Development and psychometric properties. *J Am Acad Child Adolesc Psychiatry.* 2002; 41:1061–1069. [PubMed: 12218427]
82. Russo-Ponsaran N, Yesensky J, Hessl D, et al. Feasibility, reproducibility, and clinical validity of the Pediatric Anxiety Rating Scale-Revised for fragile X syndrome. *Am J Intellect Dev Disabil.* 2013 in press.
83. Scahill L, McDougle CJ, Williams SK, et al. Children's Yale-Brown Obsessive Compulsive Scale modified for pervasive developmental disorders. *J Am Acad Child Adolesc Psychiatry.* 2006; 45:1114–23. [PubMed: 16926619]

84. Hollander E, Soorya L, Wasserman S, et al. Divalproex sodium vs. placebo in the treatment of repetitive behaviours in autism spectrum disorder. *Int J Neuropsychopharmacol.* 2006; 9:209–213. [PubMed: 16316486]
85. 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. [PubMed: 11055459]
86. Lam KS, Aman MG. The Repetitive Behavior Scale-Revised: independent validation in individuals with autism spectrum disorders. *J Autism Dev Disord.* 2007; 37:855–66. [PubMed: 17048092]
87. Brock M, Hatton D. Distinguishing features of autism in boys with fragile X syndrome. *J Intellect Disabil Res.* 2010; 54:894–905. [PubMed: 20704635]
88. Sparrow, SS.; Cicchetti, DV.; Balla, DA. *Vineland Adaptive Behavior Scales. 2.* Minneapolis, MN: Pearson Assessments; 2005.
89. Seltzer MM, Abbeduto L, Greenberg JS, et al. Biomarkers in the Study of Families of Individuals with Developmental Disabilities. *Int Rev Res Ment Retard.* 2009; 37:213–249. [PubMed: 20414357]
90. Erickson CA, Weng N, Weiler IJ, et al. Open-label riluzole in fragile X syndrome. *Brain Res.* 2011; 1380:264–270. [PubMed: 21059347]
91. Westmark C, Westmark P, O’Riordan K, et al. Reversal of fragile X phenotypes by manipulation of A PP/A levels in Fmr1 KO mice. *PLoS One.* 2011; 6:e26549. [PubMed: 22046307]
92. Gross C, Nakamoto M, Yao X, et al. Excess phosphoinositide 3-kinase subunit synthesis and activity as a novel therapeutic target in fragile X syndrome. *J Neurosci.* 2010; 30:10624–10638. [PubMed: 20702695]
93. Hoeffler CA, Sanchez E, Hagerman RJ, et al. Altered mTOR signaling and enhanced CYFIP2 expression levels in subjects with fragile X syndrome. *Genes Brain Behav.* 2012; 11:332–41. [PubMed: 22268788]
94. Hessel D, Cordeiro L, Yuhas J, et al. Prepulse inhibition in fragile X syndrome: feasibility, reliability, and implications for treatment. *Am J Med Genet.* 2008; 153B:545–553.
95. de Vrij FM, Levenga J, van der Linde HC, et al. Rescue of behavioral phenotype and neuronal protrusion morphology in Fmr1 KO mice. *Neurobiol Dis.* 2008; 31:127–132. [PubMed: 18571098]
96. Berry-Kravis E, Hessel D, Coffey S, et al. A pilot open label, single dose trial of fenobam in adults with fragile X syndrome. *J Med Genet.* 2009; 46:266–271. [PubMed: 19126569]
97. Heilman KJ, Harden E, Berry-Kravis E, et al. Autonomic regulation in fragile X syndrome. *Dev Psychobiol.* 2011; 53:785–795. [PubMed: 21547900]
98. Hall SS, Lightbody AA, Huffman LC, et al. Physiological correlates of social avoidance behavior in children and adolescents with fragile x syndrome. *J Am Acad Child Adolesc Psychiatry.* 2009; 48:320–9. [PubMed: 19182690]
99. Farzin F, Scaggs F, Hervey C, et al. Reliability of eye tracking and pupillometry measures in individuals with fragile X syndrome. *J Autism Dev Disord.* 2011; 41:1515–1522. [PubMed: 21267642]
100. Santoro MR, Bray SM, Warren ST. Molecular mechanisms of fragile X syndrome: a twenty-year perspective. *Annu Rev Pathol.* 2012; 7:219–245. [PubMed: 22017584]
101. Hazlett HC, Poe MD, Lightbody AA, et al. Teasing apart the heterogeneity of autism: Same behavior, different brains in toddlers with fragile X syndrome and autism. *J Neurodev Disord.* 2009; 1:81–90. [PubMed: 20700390]
102. Hoefft F, Lightbody AA, Hazlett HC, et al. Morphometric spatial patterns differentiating boys with fragile X syndrome, typically developing boys, and developmentally delayed boys aged 1 to 3 years. *Arch Gen Psychiatry.* 2008; 65:1087–1097. [PubMed: 18762595]
103. Hoefft F, Walter E, Lightbody AA, et al. Neuroanatomical differences in toddler boys with fragile x syndrome and idiopathic autism. *Arch Gen Psychiatry.* 2011; 68:295–305. [PubMed: 21041609]
104. Haas BW, Barnea-Goraly N, Lightbody AA, et al. Early white-matter abnormalities of the ventral frontostriatal pathway in fragile X syndrome. *Dev Med Child Neurol.* 2009; 51:593–599. [PubMed: 19416325]

105. Watson C, Hoefft F, Garrett AS, et al. Aberrant brain activation during gaze processing in boys with fragile X syndrome. *Arch Gen Psychiatry*. 2008; 65:1315–1323. [PubMed: 18981343]
106. Hallahan BP, Craig MC, Toal F, et al. In vivo brain anatomy of adult males with Fragile X syndrome: an MRI study. *Neuroimage*. 2011; 54:16–24. [PubMed: 20708694]
107. Wilson LB, Tregellas JR, Hagerman RJ, et al. A voxel-based morphometry comparison of regional gray matter between fragile X syndrome and autism. *Psychiatry Res*. 2009; 174:138–145. [PubMed: 19853418]
108. Lee AD, Loew AD, Lu A, et al. 3D pattern of brain abnormalities in Fragile X syndrome visualized using tensor-based morphometry. *Neuroimage*. 2007; 34:924–938. [PubMed: 17161622]
109. Hoefft F, Carter JC, Lightbody AA, et al. Region-specific alterations in brain development in one- to three-year-old boys with fragile X syndrome. *Proc Natl Acad Sci U S A*. 2010; 107:9335–9339. [PubMed: 20439717]
110. Bray S, Hirt M, Jo B, et al. Aberrant frontal lobe maturation in adolescents with fragile X syndrome is related to delayed cognitive maturation. *Biol Psychiatry*. 2011; 70:852–858. [PubMed: 21802660]
111. Kesler SR, Lightbody AA, Reiss AL. Cholinergic dysfunction in fragile X syndrome and potential intervention: a preliminary 1H MRS study. *Am J Med Genet*. 2009; 149A:403–407. [PubMed: 19215057]
112. Hagerman RJ, Jackson AW, Levitas A, et al. Oral folic acid versus placebo in the treatment of males with the fragile X syndrome. *Am J Med Genet*. 1986; 23:241–262. [PubMed: 3513567]
113. Hagerman RJ, Murphy MA, Wittenberger MD. A controlled trial of stimulant medication in children with the fragile X syndrome. *Am J Med Genet*. 1988; 30:377–392. [PubMed: 3052064]
114. Strom CM, Brusca RM, Pizzi WJ. Double-blind, placebo-controlled crossover study of folinic acid (Leucovorin) for the treatment of fragile X syndrome. *Am J Med Genet*. 1992; 44:676–682. [PubMed: 1481832]
115. Hagerman RJ, Miller LJ, McGrath-Clarke J, et al. Influence of stimulants on electrodermal studies in Fragile X syndrome. *Microsc Res Tech*. 2002; 57:68–73.
116. Berry-Kravis E, Krause SE, Block SS, et al. (Effect of CX516, an AMPA-modulating compound, on cognition and behavior in fragile X syndrome: a controlled trial. *J Child Adolesc Psychopharmacol*. 2006; 16:525–540. [PubMed: 17069542]
117. Torrioli MG, Vernacotola S, Peruzzi L, et al. A double-blind, parallel, multicenter comparison of L-acetylcarnitine with placebo on the attention deficit hyperactivity disorder inflexible X syndrome boys. *Am J Med Genet A*. 2008; 146:803–812. [PubMed: 18286595]
118. Erickson CA, Mullett JE, McDougale CJ. Open-Label Memantine in Fragile X Syndrome. *J Autism Dev Disord*. 2009; 39:1629–1635. [PubMed: 19609663]
119. Torrioli M, Vernacotola S, Setini C, et al. Treatment with valproic acid ameliorates ADHD symptoms in fragile X syndrome boys. *Am J Med Genet A*. 2010; 152A:1420–1427. [PubMed: 20503316]
120. Paribello C, Tao L, Folino A, et al. Open-label add-on treatment trial of minocycline in fragile X syndrome. *BMC Neurology*. 2010; 10:91. [PubMed: 20937127]
121. Erickson CA, Stigler KA, Wink LK, et al. A prospective open-label study of aripiprazole in fragile X syndrome. *Psychopharmacology (Berl)*. 2011; 216:85–90. [PubMed: 21318565]
122. Hall SS, Lightbody AA, McCarthy BE, et al. Effects of intranasal oxytocin on social anxiety in males with fragile X syndrome. *Psychoneuroendocrinology*. 2012; 37:509–518. [PubMed: 21862226]
123. Erickson, CA. Prospective open-label trial of acamprosate in youth with fragile X syndrome: a pilot study Including biomarker assays. *International Fragile X Syndrome Meeting*; July 25, 2012; Miami, Florida.

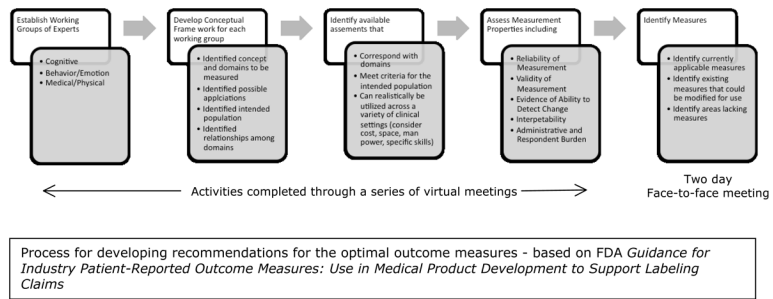


Figure 1. Process for developing recommendations for the optimal outcome measures - based on FDA *Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims*

Table 1

Outcome Measures Used in Prior Published Trials in FXS

Trial (First author, Year published (reference))	Design ¹	N ²	Measures Used ³		Cognitive	Biomarker	Measures Reliable ⁴	Measures Change- Sensitive ⁵
			Behavioral					
High-dose folate, (Hagerman, 1986 (112))	DB, PC, XO	25	CARS, AutBC		PPVT, TOLD, IQ		ND	None
Methylphenidate, (Hagerman 1988 (113))	DB, PC, XO	15	Conner Scales				ND	Conner
Folinic acid (Strom 1992 (114))	DB, PC, XO	21	CTRS, CPRS ACTeRS		VABS, PPVT		ND	None
Methylphenidate, Amphetamine salts, (Hagerman 2002 (115))	OL	19				EDR	ND	EDR
CX516, (Berry-Kravis 2006 (116))	DB, PC	48	ABC, SNAPIV, VAS, CGI, CARS, GARS, ADOS		RBANS, CELF, PLS-4, PPVT, TVPS, W-JMem, IVA, VABS		PPVT, PLS-4, RBANS (LR, LL, DS, SM subtests), W-JMem, SNAPIV, ADOS Comm, CARS, ABC Subscales, VAS Behavior	None
N-acetyl-carnitine (Torrioli 2008 (117))	DB, PC, XO	63	CGI-T, CGI-P		VABS, WISC		ND	CGI-T, CGI-T, VABS
Lithium, (Berry-Kravis 2008 (80))	OL	15	ABC, VAS, CGI		RBANS, PPVT, CFXCPT, NEPSY Tower, NVALT, Card Task, VABS	ERK, Eye Tracking, Auditory Filtering, Autonomic Testing – RSA, HR, HRV	CFXCPT (in concurrent study)	ABC, VAS, CGI, RBANS (LL), VABS-L, ERK
Fenobam, (Berry-Kravis 2009 (96))	OL	12			CFXCPT	PPI	PPI (concurrent study)	PPI
Donepezil, (Fung 2012, Kesler 2009, (73,111))	OL		ABC, CBCL		CNT, HVLT	MRS		
Mennantine, (Erickson 2009 (118))	OL	6	ABC, CGI, ADHDRS-IV, SRS				ND	None significant
Valproic acid, (Torrioli 2010 (119))	OL	10	CTRS, CPRS, CGI, SNAPIV, K-SADS-PL		VABS	FMRI re-activation	ND	CPRS-R
Minocycline, (Paribello 2010 (120))	OL	20	ABC, CGI, VAS		RBANS, PPVT, NVALT		ND	ABC, CGI, VAS, RBANS (LL, SM)
Aripiprazole, (Erickson 2010 (121))	OL	15	ABC, CGI, CY-BOCS				ND	ABC, CY-BOCS, CGI

Trial (First author, Year published (reference))	Design ¹	N ²	Measures Used ³		Cognitive	Biomarker	Measures Reliable ⁴	Measures Change- Sensitive ⁵
			Behavioral					
Ritzolz, (Erickson 2011 (90))	OL	6	CY-BOCS, CGI, ADHDRS-IV, ABC, SRS	PPVT	ERK	ND	ADHDRS-IV, ERK	
AFQ056, (Jacquemont 2011 (76))	DB, PC, XO	30	ABC, RBS, CGI, VAS, SRS	KITAP, RBANS, PPVT	ERK, Eye Tracking, PPI, fMRI methylation, FMR1 mRNA	ND	CGI, ABC, RBS, VAS (in fully methylated group)	
Oxytocin, (Hall 2012 (122))	DB, PC	8			Eye tracking, cortisol, Autonomic Testing – RSA, HR, HRV	ND	Eye tracking, cortisol	
Arbaclofen, (Berry-Kravis 2012 (72))	DB, PC, XO	62	ABC, RBS, VAS, CGI, SRS, CASI, ADHDRS-IV	RBANS, PPVT, VABS, SB5- WM	PPI, APP		ABC-SW, VABS- P/L, CGI, VAS	
Acamprosate, (Erickson 2012 (123))	OL	12	ABC, CGI, SRS, CASI, ADHDRS-IV	PPVT, VABS	APP		APP, ABC, CGI, SRS, ADHDRS-IV	

¹ DB=Double-blind, PC=Placebo-controlled, XO=Crossover, OL=Open label

² Number of subjects in trial

³ Adaptive measures are included in the Cognitive section. Abbreviations for measures as follows: ABC=Aberrant Behavior Checklist-Community Edition (SW=Social Withdrawal Subscale), ACTeRS=ADD-H Comprehensive Teacher's Rating Scales, ADHDRS-IV=ADHD Rating Scale IV, ADOS=Autism Diagnostic Observation Scale, APP=Amyloid precursor protein, AutBC=Autism Behavior Checklist, Card Task=Card Task Test of Visual Sequential Memory, CARS=Childhood Autism Rating Scale, CASI=Childhood Anxiety Sensitivity Index, CBCL=Child Behavior Checklist, CELF=Clinical Evaluation of Language Skills, CFXCPT=Carolina Project Fragile X Continuous Performance Test, CGI=Conner Global Impression, CGI-P=Conner Global Index for Parents, CGI-T=Conner Global Index for Teachers, CNT+=Contingency Naming Task, CPRS=Conner Parent Rating Scale, CTRS=Conner Teacher Rating Scale, CY-BOCS =Children's Yale-Brown Obsessive Compulsive Scale modified for pervasive developmental disorders, EDR=Electrodermal responses, ERK=ERK-related kinase activation rate, GARS=Gilliam Autism Rating Scale, HR=Heart Rate, HRV=Heart Rate Variability, HVLT=Hopkins Verbal Learning Test, IVA=Integrated Visual-Auditory CPT, KITAP=Tests of Attentional Performance for Children, K-SADS PL=Kiddie Schedule for Affective Disorder and Schizophrenia for School Age Children - Present and Lifetime Version, NVALT=Non-Verbal Associative Learning Task, PLS-4=Preschool Language Scale – Version 4, PPI=Prepulse Inhibition, PPVT=Peabody Picture Vocabulary Test, RBANS=Repeatable Battery for the Assessment of Neurological Status (LL=List Learning, SM=Story Memory, DS=Digit Span, LR=List Recall Subtests), RBS=Repetitive Behavior Scale, RSA=Respiratory Sinus Arrhythmia, SB5-WM=Stanford-Binet Version 5 Working Memory Index, SNAPIV=Swanson, Nolan, and Pelham Questionnaire for ADHD, SRS=Social Responsiveness Scale, TOLD=Test of Language Development, TVPS=Test of Visual Perceptual Skills, VABS=Vineland Adaptive Behavior Scale (P/L=Play and Leisure Subscale), VAS=Visual Analog Scale, W-JMem=Woodcock-Johnson Memory for Words

⁴ Reliability not assessed in all trials, data presented includes only those analyzed that had ICC or weighted Kappa>0.8 on test retest

⁵ Measures change-sensitive in trial are presented, it must be recognized that lack of change-sensitivity in the absence of a drug effect does not rule out change sensitivity for the measure, change sensitivity may be exaggerated or reported as present when measuring only placebo effect in open-label studies, if just a subtest was change-sensitive-subtest name given in parenthesis

Table 2

Promising Outcome Measures for the Cognitive Domain

Domain	Outcome Measure	Pros	Cons
Language	Standardized Language Sampling Procedures (e.g., narration)	Useful for wide age/ability range	Labor-intensive transcription; psychometrics not fully known for FXS
Memory	WJ Auditory Working Memory WJ Digits Reversed Corsi Blocks CANTAB Object Memory RBANS List Learning	Collectively, broad coverage of domain; good psychometric properties in general population; evidence of reliability and validity in FXS	Not all subtests appropriate for lower-functioning people
Executive Function	WJ Planning Contingency naming KITAP (4 subtests) WJ Rapid naming	Collectively, broad coverage of domain; good psychometrics in general population; evidence of reliability and validity in FXS	Not all subtests feasible for lower-functioning people
Social Cognition	Social Responsiveness Scale	Broad coverage of social behaviors; easy to administer; strong psychometrics in other populations Useful for wide age/ability range	Unknown validity for FXS
	Eye Tracking (to measure emotion processing)	Useful for wide age/ability range	Unknown validity for FXS; technology not widely available
Academic Achievement	Woodcock-Johnson Tests of Achievement-III	Wide age range; good reliability; alternate test forms	Unknown validity for FXS
	Wechsler Individual Achievement Test-II	Wide age range; validity studies with clinical populations	Unknown validity for FXS

Table 3

Promising Measures of Maladaptive Behavior

Domain	Potential Outcome Measure	Pros	Cons
Irritability/aggression	ABC-Irritability	Known sensitivity in ID/AUT. FXS psychometric data	Sensitivity in FXS unknown. Limited to lower functioning patients
Anxiety	PARS ADAMS	Good phenotype coverage; sensitive in pediatric anxiety trials, existing FXS psychometric data Validated for ID; limited FXS psychometric data	Limited data in ID and FXS sensitive unknown. Unclear validity for higher functioning patients
Hyperactivity/impulsivity	Conners/SNAP-IV	Sensitive in ADHD trials	Validity in FXS unknown
Inattention	Conners/SNAP-IV	Sensitive in ADHD trials	Validity in FXS unknown
Social reciprocity/avoidance	SRS ADAMS (social avoidance) ABC – Lethargy/withdrawal	Good coverage of domains; linked to broad autism phenotype FXS reliability and validity data FXS reliability and validity data	Validity in FXS unknown Sensitivity unknown Lethargy confounded with social items (however, see 81)
Repetitive/compulsive behaviors	Repetitive Behavior Scale-Revised	Good coverage of domains; validated in autism/ID	Limited FXS data.
Adaptive behavior	Vineland	Extensive FXS data; sensitive in double blind FXS trials	A by parental expectations and environment; wide period of observation
Self-injury	Behavior Problems Inventory	Published FXS data; severity and frequency ratings	Sensitivity to treatment unknown

Table 4

Promising Biomarkers

Domain	Potential Outcome Measure	Pros	Cons
FMR1/FMRP expression measures in blood	FMR1 methylation FMRP levels	Predicted response to AFQ056 in small study, stable, easy to store/ship Key factor in severity of molecular deficit being targeted	True predictive effect unknown, variability between assays No easy and accurate assay, no association with outcome of treatment yet, not reflect downstream events from FMRP deficit
Signaling Pathway Measures proposed as biomarkers in blood	ERK activation rate APP and metabolite levels PI3K activity mTOR activity S6 Kinase activity MMP activity	Abnormal in FXS blood, responsive to lithium tx Abnormal in FXS blood, protein regulated by FMRP, stable in serum, easy to collect, store and process samples, One study showed abnormal in FXS One study showed abnormal in FXS Levels abnormal, normalized in minocycline trial	Not responsive to tx in some trials, not stable, assay difficult, requires high blood volume Treatment response not yet shown Treatment response not yet shown Treatment response not yet shown Variable results between laboratories for levels in individuals with FXS
Biophysical Measures	Prepulse Inhibition (PPI) Eye Tracking and Pupillometry Heart Rate Variability	Good reproducibility, treatment responsive in fenobam trial Easy to do, feasible for most patients, good reproducibility Feasible for most patients	Lack of demonstration of clinical validity, expensive and difficult to standardize across sites, younger patients don't tolerate it Clinical validity not clear, not yet shown to be treatment responsive, expensive Not treatment responsive in lithium study, inconsistent methods across labs, reproducibility not defined, expensive
Neuroimaging	Structural MRI DTI MRS rsfMRI fMRI PET		Some procedures may require behavioral preparation or sedation in lower functioning affected males Functional MRI is not feasible in many affected males