Published in final edited form as:

Biol Psychiatry. 2007 February 15; 61(4): 504–511.

Developmental Disabilities Modification of Children's Global Assessment Scale (DD-CGAS)

Ann Wagner, Ph.D.¹, Luc Lecavalier, Ph.D.², L. Eugene Arnold, M.D., M.Ed.², Michael G. Aman, Ph.D.², Lawrence Scahill, MSN, Ph.D.³, Kimberly A. Stigler, M.D.⁴, Cynthia R. Johnson, Ph.D.⁵, Christopher J. McDougle, M.D.⁴, and Benedetto Vitiello, M.D.¹

¹ National Institute of Mental Health, National Institutes of Health

² Ohio State University

³ Yale University

⁴ Indiana University School of Medicine

⁵ University of Pittsburgh School of Medicine

Abstract

Background—Interventions for pervasive developmental disorders (PDD) aim to alleviate symptoms and improve functioning. To measure global functioning in treatment studies, the Children's Global Assessment Scale was modified and psychometric properties of the revised version (DD-CGAS) were assessed in children with PDD.

Methods—Developmental disabilities-relevant descriptors were developed for the DD-CGAS and administration procedures were established to enhance rater consistency. Ratings of clinical case vignettes were used to assess inter-rater reliability and temporal stability. Validity was assessed by correlating the DD-CGAS with measures of functioning and symptoms in 83 youngsters with PDD. Sensitivity to change was assessed by comparing change from baseline to post-treatment with change on the Aberrant Behavior Checklist – Irritability and Clinical Global Impressions–Improvement subscale scores in a subset of 14 children.

Results—Inter-rater reliability (ICC=.79) and temporal stability (average ICC = .86) were excellent. DD-CGAS scores correlated with measures of functioning and symptoms with moderate to large effect sizes. Changes on the DD-CGAS correlated with changes on the ABC-I ($r=-.71$) and CGI-I ($r=-.52$). The pre-post DD-CGAS change had an effect size of .72.

Conclusions—The DD-CGAS is a reliable instrument with apparent convergent validity for measuring global functioning of children with PDD in treatment studies.

Keywords

autism; pervasive developmental disorder; functioning; assessment; children; psychometrics

Correspondence to: Dr. A. Wagner, NIMH Room 6184, 6001 Executive Blvd., Bethesda, MD 20982-9633. Tel.: 301/443-3633, Fax: 301/480-4415, e-mail: awagner@mail.nih.gov

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Introduction

Functional impairment is a critical aspect of mental illness. It is the functional impact of psychiatric/behavioral symptoms that often prompts clinical referral and treatment. The efficacy of treatment is traditionally established based on symptomatic improvement, but this is a limited perspective in need of validation by demonstrating parallel functional improvement. Documenting treatment effects on functioning is especially relevant to children with Autistic Disorder (autism) and other Pervasive Developmental Disorders (PDD) (Arnold et al 2000). There are currently no curative treatments for the core deficits in social interaction, communication, and repetitive and/or rigid behaviors (American Psychiatric Association 1994). However, there is evidence that both behavioral and pharmacological interventions can significantly ameliorate core symptoms, as well as improve adaptive skills and decrease commonly associated behavior problems such as aggression and hyperactivity (Horner et al 2002; Lovaas 1987; McEachin et al 1993; National Research Council 2001; Sallows & Graupner, 2005; RUPP Autism Group 2002 and 2005).

In clinical trials, the assessment of treatment effects on functioning of children with PDD is hampered by the lack of reliable, sensitive, and easy to administer global rating instruments. Several scales exist for rating level of functioning in adults with mood, anxiety, and psychotic disorders (Endicott et al, 1976; Endicott et al 1997; Weissman et al 2001). The *Children's Global Assessment Scale* (CGAS) (Schaffer et al 1983) is a modification of the *Global Assessment Scale* (GAS) for adults (Endicott et al, 1976). It is commonly used for rating functioning in children, and was found to be sensitive to treatment effects in adolescents with depression (Mufson et al 2004). The descriptors of the CGAS scores, however, are not all relevant to PDD and cannot be easily applied to children with these disorders who typically follow abnormal developmental trajectories and present with severe impairments in specific areas of functioning. Intellectual functioning can range from profound mental retardation to the superior range, and frequently there are discrepancies between intellectual and adaptive skills, usually with adaptive skills lagging behind mental age (Bolte & Poutska 2002; Schatz & Hamdan-Allen 1995; Stone et al, 1999). An instrument to assess global functioning would need to accommodate a wide range of functioning with substantial variability both between and within subjects, and integrate information about multiple domains of functioning.

Although instruments such as the Vineland Adaptive Behavior Scales (VABS; Sparrow et al 1984; Volkmar et al 1993) and the Assessment of Basic Language and Learning Skills (ABLLS; Partington and Sundberg 1998) can be used to measure specific areas of adaptive behavior in children with PDD, their sensitivity to differential treatment effects in clinical trials has not been established (Smith et al., 2000). These instruments are lengthy to administer and restricted to specific domains of functioning.

In spite of the considerable individual variability in level of functioning across specific domains, global ratings of functioning are useful summary measures that are clinically meaningful, incorporate all available sources of information, and help gauge the overall therapeutic value of interventions. There is also evidence suggesting that global ratings can be more sensitive to change during acute treatment than scores on itemized symptom rating scales (Lehmann, 1984; Endicott et al, 1976). In fact, by integrating information from various sources about a subject's functioning, global ratings provide a more comprehensive view than scores based on specific scales and a single informant can offer. The Clinical Global Impressions Scale (CGI; Guy 1976) is often used in clinical trials, including with children with PDD, as a global measure of severity of illness and improvement, but is generally focused on symptoms-- sometimes a specific cluster of symptoms-- rather than on functional impairment.

Given the absence of a rating instrument that yields a quantitative measure of global functioning for use in clinical trials involving children with developmental disabilities, the CGAS was modified by adapting the anchor points and the administration procedure to the characteristics of children with developmental disabilities including PDD. This report describes the *Developmental Disability-Child Global Assessment Scale* (DD-CGAS) and presents data on its inter-rater reliability, temporal stability, convergent validity, and sensitivity to change during treatment when applied to a population of children with PDD.

Methods and Materials

Description of the DD-CGAS

The DD-CGAS was modified from the CGAS (Shaffer et al 1983). It is a clinician-rated scale yielding a single score of global functioning of a child (here defined as a subject under 18 years of age) with a developmental disability relative to his or her typically developing same-age peers. The rating reflects typical functioning of the child during a particular time period, usually the week prior to the evaluation. The rating is intended to be a global rating based on all available sources of information and across all domains of functioning, including self care, communication, social behavior, and school/academic functioning. The rating is not meant to be dependent on the particular diagnosis, perceived cause of dysfunction (e.g., cognitive or physical limitation, environmental constraints, behavioral disturbance), or type and severity of symptoms.

Maintaining the overall structure of the original GAS and CGAS, the DD-CGAS is a dimensional scale with scores ranging from 1 to 100, where 1 represents the most impaired functioning and 100, superior functioning. Each decile (e.g., 1–10, 11–20) has a descriptive header (e.g., “*Moderate impairment in functioning in most domains*”) and examples of behaviors and types of environmental accommodations that might be seen at that level of functioning (see Figure 1). Scores above 70 on the DD-CGAS indicate functioning within the range of typically developing children of the same age as the child being rated. Since children with developmental disabilities must have, by definition, significant functional impairment, one would seldom give ratings above 70 in this population. However, children with mild disabilities may improve with treatment to a degree that they are functioning within the normal range. Furthermore, because this instrument is intended to be useful for a variety of types of research and with a range of developmental disabilities and control groups, an instrument capturing the full range of functioning was desired.

Because of the critical role that clinical judgment has on global ratings, a specific procedure was devised to standardize the approach of scoring the DD-CGAS in order to increase reliability. To this end, a scoring grid (Figure 2) was developed that assigns a level of impairment (none, slight, moderate, severe, extreme) to four key domains of functioning (self care, communication, social behavior, and school/academic). The rater first determines the level of impairment for each domain, taking into consideration the child’s behavior, consistency across settings (e.g., home, school, and community), level of environmental adaptation needed to support the child, and level of supervision required. Then the rater chooses the interval heading that best describes the levels of functioning across the domains (e.g., 50–41: “*Moderate impairment in functioning in most domains and severe impairment in at least one domain*”). The examples within the interval headings are used to confirm the description of the child’s functioning, although no child will be perfectly described by these hypothetical descriptions.

When the “best fit” interval has been determined, the rater considers the adjacent intervals in order to assign a specific rating. For example, if the child fits best into “60–51: *Moderate impairment in functioning in most areas*” but has some similarity to 41–50, the rater applies a

number in the lower half of the range (i.e., 54–51). Conversely, if the child fits best in 60–51 but has some strengths consistent with the next higher category, the rater would apply a number in the top half of the category (i.e., 60–56).

All available sources of information should be used to make the rating. This might include direct observation, caregiver reports, and results of standardized tests. Whatever the source, the rater needs a good description of the functioning in key domains and across multiple settings. The scale then allows the rater to synthesize all available information into a single index of functioning. The amount of time to gather relevant information will vary with the situation in which the instrument is being used. Once that information is gathered, it takes between 5 and 10 minutes to make the initial rating. Re-rating the same child usually takes less time.

Inter-Rater Reliability and Temporal Stability

Written vignettes were derived from sixteen clinical cases reflecting a range of functioning among children with PDD. Vignettes described children between 4 and 14 years of age, inclusive. Nine (56%) of the vignettes described boys. IQ scores ranged from 20 to 98. The vignettes (3–5 pages in length) included age and sex of the child, as well as extensive behavioral descriptions of behavior and functioning in the following areas: self-care skills (including eating/feeding, dressing/undressing, sleeping, toileting, performing daily routines), communication (including verbal language skills, social communication, nonverbal communication, reading/writing), social behavior (including family relationships, peer relationships, and level of appropriate/inappropriate social behavior), and school functioning (including placement, academic achievement, and adaptive behavior in school). Vignettes also included a description of consistency/inconsistency across settings, level of environmental adaptations needed, and level of supervision required. Gold standard scores for these Reliability Vignettes were derived from the average of the six developers' ratings on each vignette. Gold standard ratings of the vignettes ranged from 24 to 73.

Thirteen clinicians independently rated the clinical vignettes to assess inter-rater reliability. The raters varied in level of training and experience, but all were involved in multi-site clinical research with children with PDD. They had familiarized themselves with the DD-CGAS scoring and had discussed and reviewed together six or more vignettes for training purposes. These raters were located at five different sites, including Indiana University, National Institute of Mental Health, Ohio State University, the University of Pittsburgh, and Yale University. Eight of the thirteen clinicians were available to rate the clinical vignettes again after 3 – 7 months, for an assessment of temporal stability. They had not been told that they would be asked to complete the ratings a second time.

Validity and Sensitivity to Change

Procedure—The DD-CGAS was included in an ongoing RUPP Autism Network intervention study. Independent evaluators for the study were certified to administer the DD-CGAS by teleconference training sessions that included rating the clinical vignettes described above. The raters independently rated six of the Reliability Vignettes that had been assigned gold standard ratings by the developers. An individual was considered certified if he or she was within 10 points of the gold standard on 80% of the vignettes. If a rater failed to become certified, he or she had another training session and then rated another set of six vignettes. A third trial of four ratings was available if needed. All raters but one achieved certification within two trials; the seventh rater achieved certification on the third trial.

The intervention study consisted of a small pilot study and a randomized clinical trial. The DD-CGAS was administered by an independent evaluator according to the rating instructions in

Figures 1 and 2, using all available clinical and test data. Subjects from both the pilot and randomized trial contributed baseline test scores for assessing the DD-CGAS's validity. Post-intervention data (after 24 weeks of intervention) was available from a subset of the pilot subjects for a preliminary evaluation of the DD-CGAS's sensitivity to change.

Subjects—The pilot study and randomized trial protocols were approved by the following institutional review boards (IRB): Ohio State University Behavioral and Social Sciences IRB, the Yale IRB, and the Indiana University/Perdue University at Indianapolis and Clarion IRB. Pittsburgh University participated in the pilot study only and that protocol was approved by the University of Pittsburgh IRB. Informed consent for human investigation was obtained from parents of the participants.

A total of eighty-three subjects contributed baseline scores to assess concurrent validity. Seventeen were from the pilot study and 66 were from an ongoing randomized clinical trial. Subjects had an IQ of ≥ 35 or a mental age ≥ 18 months. The average age was 7.62 years ($SD=2.54$ years; range 4.09 – 13.81 years). Sixty-five subjects (78%) were boys. Diagnoses were as follows: Autistic Disorder, 56; Pervasive Developmental Disorder, Not Otherwise Specified (PDDNOS), 21; and Asperger's Disorder, 6. Diagnoses were established by clinical assessment and corroborated with the Autism Diagnostic Interview–Revised (Lord et al, 1994). The DD-CGAS scores at baseline ranged from 11 – 68. Table 1 shows subject characteristics.

Post-intervention data were available for the subset of fourteen pilot study subjects. The average age of this group was 8.33 years ($SD=2.75$ years; range 4.12 – 13.73). Eleven (79%) were boys and diagnoses were as follows: Autistic Disorder, 9; PDDNOS, 3; and Asperger's Disorder, 2.

Instruments—*Vineland Adaptive Behavior Scale – Survey Form* (VABS; Sparrow et al., 1984) is a standardized measure of adaptive functioning based on parent interview. The Adaptive Behavior Composite is a total score with a mean of 100 and SD of 15. Higher scores indicate more mature adaptive functioning.

Assessment of Basic Language and Learning Skills (ABLLS; Partington & Sundberg, 1998) is a criterion referenced measure of adaptive skills. It contains 26 subscales. Raw scores from five subscales (dressing/clothing; eating/meal preparation; grooming; toileting; household chores/tasks) were chosen because of their relevance to the interventions being tested and were summed to provide a composite score. Higher scores indicate more mature adaptive skills.

Stanford-Binet Intelligence Scale Fifth Edition (SB5; Roid, 2003) is a standardized individual measure of intellectual functioning that covers an age range from 2 years to adulthood. The test yields standardized IQ scores with a mean of 100 and SD of 15.

Leiter International Performance Scale – Revised (Leiter-R; Roid & Miller, 1997) is a nonverbal test of intelligence for children and adolescents between the ages of 2 and 20 years. The test yields a composite score with a mean of 100 and a SD of 15.

Aberrant Behavior Checklist (ABC; Aman et. al., 1985a,1985b) is a 58-item informant-based scale comprising five subscales. The 16-item Irritability subscale was used in the current study because of its relevance to the intervention. Items are rated on a four-point scale; higher scores indicate more severe problem behavior.

Children's Yale-Brown Obsessive Compulsive Scale – PDD (CY-BOCS–PDD) is a semi-structured, clinician-rated instrument designed to measure the current severity of repetitive behavior in children with PDD (Scahill et. al., 2006). It is a modified version of the CY-BOCS

(Scahill et al., 1997). The CY-BOCS-PDD was administered as a semistructured interview with the parents. The total score was used in the current study. Higher scores indicate more severe symptomatology.

Home Situations Questionnaire (HSQ; Barkley, 1997) is a 25-item informant based rating scale. Parents endorse the number of real-life settings in which their child is likely to be noncompliant, and rate the severity of noncompliance. The instrument was modified for this study by adding some items that reflected the types of situations that often pose challenges for children with PDD, and the instructions were altered. The mean severity score is a summary score of noncompliance; higher scores indicate greater noncompliance.

Autism Diagnostic Interview-Revised (Lord et al., 1994) is a semi-structured interview that measures the core symptoms of autism. Domain scores are derived for Social, Communication (either Verbal or Nonverbal), and Repetitive Behaviors. Although higher scores indicate greater impairment, it is designed as a categorical measure and provides an algorithm for a diagnosis of Autistic Disorder.

Clinical Global Impressions Scale (CGI; Guy, 1976) is a standard measure for making global assessments of illness. The CGI yields a Severity rating (CGI-S), an assessment of the current severity of symptoms, and an Improvement rating (CGI-I), a comparison of the individual's baseline condition to the current severity of symptoms. Ratings are made by a clinician on a seven point Likert scale using all available information about the individual's symptoms. Lower scores indicate less severe illness on the Severity scale and greater improvement on the Improvement scale. For this study, the CGI-S anchor points were modified so that "uncomplicated autism" (without accompanying behavioral or emotional problems) was assigned a score of 3 (mildly ill) (Arnold et al., 2000).

Data Analysis

Inter-rater reliability and temporal stability—Intraclass correlation coefficients (ICC) were computed to assess inter-rater reliability using thirteen independent raters' initial scores on Reliability Vignettes. ICCs were also computed on the scores of the Reliability Vignettes to assess temporal stability.

Convergent validity—Baseline DD-CGAS scores were available from 83 study subjects. Convergent validity was assessed with Pearson correlation coefficients between the DD-CGAS and other baseline clinical measures. To limit the number of correlations and reduce the probability of type I error, total or composite scores were utilized when available. The Adaptive Behavior Composite score of the VABS and IQ were treated as ordinal variables. Only algorithm items from the ADI-R were used. Because some subjects had missing data, not all correlations are based on the same sample size. Because of the range of intellectual and language skills, not all subjects were administered the same IQ test. Of the IQ measures, only the SB-5 and Leiter-R were used with a sufficient number of subjects to warrant meaningful analyses. Given the descriptive nature of the analyses, and the primary interest in the value of the correlation coefficient, we did not correct for multiple comparisons and set the alpha value at .05. Given the small sample sizes in some correlational analyses, associations should also be interpreted in terms of effect sizes. According to Cohen's (1988) guidelines, correlations of $\geq .10$ represent small effects; $\geq .30$, moderate effects; and $\geq .50$, large effects.

Sensitivity to change—Fourteen pilot subjects contributed baseline and post-intervention scores. The DD-CGAS's sensitivity to change during treatment was assessed by correlating pre-post changes on the DD-CGAS with changes on the ABC-Irritability and CGI-I scores. Pooled standard deviations were used to calculate effect sizes from baseline to post-treatment.

Results

Inter-rater Reliability and Temporal Stability

The ICC for the thirteen raters across all sixteen vignettes was .79 ($p < .001$). The ICCs between test and re-test ratings for all eight raters varied from .66 to .97 and averaged .86. All ICCs were significant at the $p < .001$ level.

Convergent Validity

Correlations between the DD-CGAS and other measures are presented in Table 2. With alpha value set at .05, the DD-CGAS was significantly and positively correlated with measures of functioning: the VABS Composite ($r = 0.50, p < .001$), ABLLS total score ($r = 0.52, p < .001$), SB-5 Composite Score ($r = 0.47, p < .001$), and Leiter-R Full Scale IQ ($r = 0.49, p < .001$). Of the measures of symptom severity, the DD-CGAS was significantly and negatively correlated with the ABC-I ($r = -0.30, p = .006$), the CY-BOCS total score ($r = -0.29, p = .008$), mean HSQ severity score ($r = -0.26, p = .016$), ADI-R Social Domain ($r = -0.30, p = .005$), ADI-R Communication Domain – Nonverbal ($r = -0.45, p = .037$), and CGI-S ($r = -0.48, p < .001$). It did not correlate significantly with the ADI-R Communication Domain – Verbal or the ADI-R Repetitive Behavior Domain.

Measuring Change

The correlation between change in DD-CGAS scores and change on the ABC Irritability subscale was $-.71$ ($n = 13, p < .01$). The correlation between change in DD-CGAS scores and CGI-Improvement at week 24 was $-.52$ ($n = 14, p = .05$). The Mean DD-CGAS score at baseline was 46.2 ($SD = 12.1$), and 54.1 ($SD = 9.7$) at post-treatment (paired t value = $-4.3; p = .001$). The mean DD-CGAS change score was 7.9 points (95% CI = 4.24 – 11.56). The effect size for the DD-CGAS was .72 ($n = 14$). The effect size for the ABC Irritability scale was .75 ($n = 13$).

Discussion

The DD-CGAS is a clinician rating of global functioning for children with PDD. Specifically designed to accommodate a wide range of functioning, with both inter- and intra-subject variability in degree and type of impairment, it is accompanied by instructions and a scoring grid to assist with rating. The DD-CGAS was found to have excellent inter-rater reliability and temporal stability over an interval of several months when raters based their scores on clinical vignettes. Reliability was obtained with a diverse group of raters, in terms of background and level of expertise, from multiple research institutions. When used in an ongoing intervention study and administered by trained raters, the scale converged well with other measures of functioning and symptoms. Preliminary data from an uncontrolled pilot study suggest that the instrument may be sensitive to clinical change during treatment.

The heterogeneity of the PDD population poses challenges to assigning global ratings. The consistency between raters was greatly enhanced with the use of the specific scoring instructions and the accompanying scoring grid. Training procedures that included practice scoring of clinical case vignettes were also probably necessary for obtaining these results. Without these procedures, the reliability of the instrument is likely to be less optimal.

Correlations between the DD-CGAS and other measures of functioning and symptoms were moderate (Cohen, 1988; Kraemer, 2005), within the range one would expect when instruments measure different but related constructs. Correlations with measures of adaptive skills and IQ suggested about 25% shared variance. Some overlap with IQ is expected, since IQ imposes limits on optimal functioning. Overlap with measures of adaptive skills is also expected, but in addition to skills measured by the VABS and the ABLLS, raters take into account the degree

of environmental accommodation or support necessary to achieve a certain level of functioning. Since environmental accommodations such as 1:1 assistance in school, alternative and augmentative communication systems, and self-contained classrooms are common elements of intervention programs, it is important that a rating system take into account the level of support needed for a child to function optimally.

Most measures of symptoms were moderately correlated with the DD-CGAS. While the DD-CGAS does not measure symptoms per se, one expects that symptoms will have an impact on functioning. It appears that the DD-CGAS is sensitive to the effect of core social and communication deficits, irritability, obsessive-compulsive symptoms, and noncompliance on functioning. Two domain scores from the ADI-R did not correlate significantly with the DD-CGAS: the ADI-R Communication Domain–Verbal, and the ADI-R Repetitive Behavior Domain. The impact of communication deficits on functioning may be less with verbal children with PDD than with nonverbal children (who are likely to have cognitive limitations as well) and too subtle to be reflected in the DD-CGAS ratings. The lack of a significant correlation with the ADI-R Repetitive Domain may indicate that the presence of repetitive behaviors, narrow interests, and other symptoms captured by this subscale did not have a strong impact on rating of functional adaptation. This finding needs to be interpreted cautiously, however, as the ADI-R algorithm subscales were not constructed as an interval scale. The CYBOCS may better capture functional impairment due to excessive rigidity and repetitive behavior. Nevertheless, the small but significant correlation with the CY-BOCS suggests that the instruments are measuring different constructs.

The DD-CGAS and the CGI–S, both global clinician ratings, were also moderately correlated. Given shared variance of about 25%, the two measures are not redundant, suggesting that they are measuring different constructs, as intended.

Preliminary evidence from a subset of subjects suggested that the DD-CGAS may be sensitive to treatment effects, although the small sample size necessitates caution in interpreting the results of this uncontrolled trial. It is important to note that in addition to the small number of subjects, this was not a randomized trial, so one cannot conclude that the change was related to the treatment. However, the effect size for the DD-CGAS was medium to large and similar to the effect size of the ABC-I, which has been shown to be sensitive to treatment effects. Additionally, change in DD-CGAS scores were strongly correlated with the CGI–I. Still, one cannot rule out general bias toward assigning better scores on all instruments after participation in an intervention. Use of the DD-CGAS in a randomized, controlled trial is needed to determine its sensitivity to change and differential treatment effects.

Limitations

This study had several limitations. The measurements were made by raters who were involved in clinical research with PDD at academic sites. Extrapolation of the results to usual practice settings should be made with caution. Reliability was estimated with ratings of clinical vignettes. One cannot assume that reliability would be the same if the DD-CGAS were administered by clinicians independently assessing children and interviewing parents. Further assessment of reliability using methods that more closely resemble its intended use is needed. Some insignificant correlations might reach statistical significance with a larger sample (false negative in this report). On the other hand, our analyses did not correct for a high number of correlations and one or more might have reached significance by chance (false positive). The actual *p* values are presented so the readers can draw their own conclusions. The sample size used here was not large enough to evaluate whether subject characteristics, such as IQ and age, impact the psychometric properties of the instrument. The age range was somewhat restricted. The utility of the DD-CGAS for use with young preschoolers and older adolescents has not been demonstrated. Sensitivity to change was measured in an open-label pre-post fashion rather

than in a controlled clinical trial. Thus, we cannot rule out general bias toward assigning better scores after a period of intervention, nor can we conclude at this time that it is sensitive to differential treatment effects.

In summary, with appropriate training the DD-CGAS is a reliable assessment of global functioning that was designed to accommodate the heterogeneity found in PDD. It incorporates multiple sources of information and is quick to administer once the information is accumulated. It appears suitable for use in clinical trials with children with PDD.

The opinions and assertions contained in this report are the private views of the authors and are not to be construed as reflecting the views of the Department of Health and Human Services, the National Institutes of Health, or the National Institute of Mental Health.

Acknowledgements

This study was part of research activities of the Research Units on Pediatric Psychopharmacology (RUPP) Autism Network and funded by the following cooperative agreement grants from the National Institute of Mental Health: U10MH66768 (P.I.: M. Aman), U10MH66766 (P.I.: C. McDougle), and U10MH66764 (P.I.: L. Scahill). Janssen Pharmaceutica provided medication for the clinical trial from which some of this data were derived. Drs. Aman, Scahill, and Stigler have affiliations with Janssen Pharmaceutica. We thank Louise Ritz, Stacie Trollinger, Dawn Bozzolo, Lindsay Crowl, Kathy Koenig, Arlene Kohn, Mary Ellen Pachler, Krista Pappas, and Jennifer Wilkerson for assistance with this project.

References

- Aman MG, Novotny S, Samango-Sprouse C, Lecavalier L, Leonard E, Gadow KD, et al. Outcome Measures for Clinical Drug Trials in Autism. *CNS Spectrums* 2004;9:36–47. [PubMed: 14999174]
- Aman MG, Singh NN, Stewart AW, Field CJ. The Aberrant Behavior Checklist: A Behavior Rating Scale for the Assessment of Treatment Effects. *American Journal of Mental Deficiency* 1985a;89:485–491. [PubMed: 3993694]
- Aman MG, Singh NN, Stewart AW, Field CJ. Psychometric characteristics of the Aberrant Behavior Checklist. *American Journal of Mental Deficiency* 1985b;89:492–502. [PubMed: 3158201]
- American Psychiatric Association. *DSM-IV*. 4. Washington, DC: American Psychiatric Association; 1994. Diagnostic and Statistical Manual of Mental Disorders.
- Arnold LE, Aman MG, Martin A, Collier-Crespin A, Vitiello B, Tierney E, et al. Assessment in Multisite Randomized Clinical Trials of Patients with Autistic Disorder. *J Autism Dev Disord* 2000;30:99–111. [PubMed: 10832774]
- Barkley, RA. *Defiant Children HSQ*. New York: Guilford Publishing; 1997.
- Bolte S, Poutska F. The Relation Between General Cognitive Level and Adaptive Behavior Domains in Individuals with Autism With and Without Co-morbid Mental Retardation. *Child Psychiatry Hum Dev* 2002;33:165–172. [PubMed: 12462353]
- Charman T, Howlin P, Berry B, Prince E. Measuring Developmental Progress of Children with Autism Spectrum Disorder on School Entry Using Parent Report. *Autism* 2004;8:89–100. [PubMed: 15070549]
- Cohen, J. *Statistical power analysis for the behavioral sciences*. 2. Hillsdale, NJ: Erlbaum; 1988.
- Eikeseth S, Smith T, Jahr E, Eldevik S. Intensive Behavioral Treatment at School for 4–7 Year Old Children with Autism. A One-Year Comparison Controlled Study. *Behav Modif* 2002;26:49–68. [PubMed: 11799654]
- Endicott J, Nee J. Endicott Work Productivity Scale (EXPS): A New Measure to Assess Treatment Effects. *Psychopharmacol Bull* 1997;33:13–16. [PubMed: 9133746]
- Endicott J, Spitzer RL, Fleiss JL, Cohen J. The Global Assessment Scale: A Procedure for Measuring Overall Severity of Psychiatric Disturbance. *Arch Gen Psychiatry* 1976;33:766–771. [PubMed: 938196]
- Guy, W. Publication No. (ADM) 91–338. U.S. Dept. of Health and Human Services; Rockville, Maryland: 1976. ECDEU Assessment Manual for Psychopharmacology, Revised.

- Horner RH, Carr EG, Strain PS, Todd AW, Reed HK. Problem Behavior Interventions for Young Children with Autism: A Research Synthesis. *J Autism Dev Disord* 2002;32:423–446. [PubMed: 12463518]
- Kraemer HC, Morgan GA, Leech NL, Gliner JA, Vaske JJ, Harmon RJ. Measures of Clinical Significance. *J Am Acad Child Adolesc Psychiatry* 2005;42:1524–1529. [PubMed: 14627890]
- Lehmann E. Practicable and Valid Approaches to Evaluate the Efficacy of Nootropic Drugs by Means of Rating Scales. *Pharmacopsychiatry* 1984;17:71–75. [PubMed: 6377335]
- Lord C, Rutter M, LeCouteur A. Autism Diagnostic Interview—Revised: A Revised Version of a Diagnostic Interview for Caregivers of Individuals with Possible Pervasive Developmental Disorders. *J Autism Dev Disord* 1994;24:659–685. [PubMed: 7814313]
- Lovaas OI. Behavioral Treatment and Normal Educational and Intellectual Functioning in Young Autistic Children. *J Consult Clin Psycho* 1987;55:3–9.
- McEachin JJ, Smith T, Lovaas OI. Long-term outcome for children with autism who received early intensive behavioral treatment. *Am J Ment Retard* 1993;97(4):359–372. [PubMed: 8427693]
- Mufson L, Dorta KP, Wickramaratne P, Nomura Y, Olfson M, Weissman MM. A Randomized Effectiveness Trial of Interpersonal Psychotherapy for Depressed Adolescents. *Arch Gen Psychiatry* 2004;61:577–84. [PubMed: 15184237]
- National Research Council. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press; 2001. *Educating Children with Autism*.
- Partington, JW.; Sundberg, ML. *The Assessment of Basic Language and Learning Skill*. Pleasant Hills, CA: Behavior Analysts, Inc; 1998.
- Research Units on Pediatric Psychopharmacology (RUPP) Autism Network. Risperidone in Children with Autism for Serious Behavioral Problems. *N Engl J Med* 2002;347(5):314–321. [PubMed: 12151468]
- Research Units on Pediatric Psychopharmacology (RUPP) Autism Network. A Randomized Controlled Crossover Trial of Methylphenidate in Pervasive Developmental Disorders with Hyperactivity. *Arch Gen Psychiatry* 2005;62:1266–1274. [PubMed: 16275814]
- Roid, GH. *Stanford-Binet Intelligence Scales (SB5)*. 5. Chicago IL: Riverside Publishing; 2003.
- Roid, GH.; Miller, LJ. *Leiter International Performance Scale – Revised*. Wood Dale, IL: Stoelting; 1997.
- Sallows GO, Graupner TD. Intensive Behavioral Treatment for Children with Autism: Four-year Outcome and Predictors. *Am J Ment Retard* 2005;110:417–438. [PubMed: 16212446]
- Scahill L, Riddle MA, McSwiggin-Hardin M, Ort SI, King RA, Goodman WK, et al. Children's Yale-Brown Obsessive Compulsive Scale: Reliability and Validity. *J Am Acad Child Adolesc Psychiatry* 1997;36:844–852. [PubMed: 9183141]
- Scahill L, McDougle CJ, Williams SK, Dimitropoulos A, Aman MG, McCracken JT, et al. Children's Yale-Brown Obsessive Compulsive Scale modified for pervasive developmental disorders. *J Am Acad Child Psy* 2006;45 (9):1114–1123.
- Schatz, Hamden-Allen. Effects of Age and IQ on Adaptive Behavior Domains for Children with Autism. *J Autism Dev Disord* 1995;25:51–60. [PubMed: 7608034]
- Shaffer D, Gould MS, Brasic J, Ambrosini P, Fisher P, Bird H, et al. A Children's Global Assessment Scale (CGAS). *Arch Gen Psychiatry* 1983;40:1228–31. [PubMed: 6639293]
- Sparrow, S.; Balla, D.; Cicchetti, D. *The Vineland Adaptive Behavior Scales*. Circle Pines, MN: American Guidance Service; 1984.
- Stone WL, Ousley OY, Hepburn SL, Hogan KL, Brown CS. Patterns of Adaptive Behavior in Very Young Children with Autism. *Am J Ment Retard* 1999;104:187–199. [PubMed: 10207581]
- Volkmar FR, Carter A, Sparrow SS, Cicchetti DV. Quantifying Social Development in Autism. *J Am Acad Child Adolesc Psychiatry* 1993;32:627–632. [PubMed: 7684364]
- Weissman MM, Olfson M, Gameraoff MJ, Feder A, Fuentes M. A Comparison of Three Scales for Assessing Social Functioning in Primary Care. *Am J Psychiatry* 2001;158:460–466. [PubMed: 11229989]
- Williams SK, Scahill L, Vitiello B, Aman MG, Arnold LE, McDougle CJ, et al. Risperidone and Adaptive Behavior in Children with Autism. *J Am Acad Child Adolesc Psychiatry*. in press

Developmental Disability-Children's Global Assessment Scale

Review the subject's performance across the main domains of functioning [a] **self care, eating, dressing, sleeping**; b) **communication**; c) **social behavior**; and d) **academic performance**] and settings [home, school, and community]. Score overall level of functioning by selecting the heading that describes functioning relative to typically developing child of the same age. Use intermediary levels (e.g., 35, 58, 62) as needed. Rate actual functioning regardless of treatment or prognosis. Focus on functional interference of psychopathology rather than symptoms per se. The descriptors provided below are only illustrative and are not required for a particular rating (see Instructions for scoring details).

Specified Time Period:

- 100-91 Superior functioning. Superior functioning within family, school, with peers. Superior accomplishments relative to age peers (e.g., high achievement in Scouts). School-age child doing well academically. Independently performs daily activities and self-care appropriate for age.
- 90-81 Adequate functioning in all areas: home, school, and peers; brief disturbances of behavior or emotional distress in response to life stresses (e.g., unanticipated changes in daily routine or physical environment), but no interference with functioning. Adaptive skills at age level in all domains.
- 80-71 Slight impairment in functioning. Most daily living activities at age level, but may need prompts and structure to accomplish. Minor changes in daily routine or environment may cause transient decrease in functioning. Social interactions may be one-sided and activity-based rather than intimacy-based. May appear immature, but not deviant. Language generally age-appropriate but conversations may be one-sided and/or focused on preoccupations.
- 70-61 Slight impairment in functioning and moderate impairment in at least one domain. Social deficits apparent in most situations. Learns appropriate social skills, but inflexibly and unable to generalize. Adaptive/self-help skills immature in most areas. Behavior noticeably unusual in some situations (e.g., social groups, unstructured settings) affecting social acceptance, and may restrict participation in age-normative activities in one or two domains or in a specific setting.
- 60-51 Moderate impairment in functioning in most domains. Needs considerable structure and supervision for daily routines. Daily living/adaptive skills are below age level. Communicates needs, responds to basic requests (verbally or nonverbally). Verbal language, if present, is inflexible and delayed. Social deficits and/or unusual behaviors are apparent in most settings and contribute to functioning below age expectation.
- 50-41 Moderate impairment in functioning in most domains and severe impairment in at least one domain (e.g., daily living or communication). Social overtures and/or responses are markedly absent or inappropriate. Daily living skills significantly delayed (e.g., dressing, bathing, eating). Stereotypic and/or other persistent unusual behaviors are noticeable to a casual observer and impede functioning.
- 40-31 Severe impairment in functioning in some domains. Rudimentary instrumental (not social) communication skills. Repetitive behaviors that interfere with adaptive functioning. Marked social withdrawal in most situations. Adaptive behavior significantly impaired. Significant environmental accommodations are needed in some domains. Very immature adaptive and self-care skills in at least two domains.
- 30-21 Severe impairment in all domains and settings, (e.g., home and school). Markedly withdrawn and isolated behavior. Requires extensive environmental accommodations (e.g., 1:1 supervision for behavior, locking cabinets, removing breakable objects from bedroom). Dependent in all aspects of daily living (e.g., dressing, bathing, toileting) beyond age expectation. May exhibit disturbance of basic regulatory process (e.g., sleeping, feeding).
- 20-11 Extreme impairment in at least one domain. Needs constant supervision; or extensive environmental accommodations for safety or for basic care (e.g., feeding, toileting). May need residential placement. Does not communicate basic needs. Does not interact with others. Marked disturbance of basic regulatory processes (e.g., sleeping, feeding).
- 10-1 Extreme and pervasive impairment. Poses danger to self or others. Needs intensive constant supervision (e.g., 24-hr care outside of the home) for safety or total dependence in basic self-help skills. Marked disturbance of basic regulatory processes. Needs require specialized care (e.g., behavior management or medical care) beyond what can be provided at home and outpatient support services.

The DD-CGAS was adapted from the Children's Global Assessment Scale (CGAS; Shaffer et al, 1983) and the Global Assessment Scale (GAS; Endicott et al, 1976)

Figure 1.
DD-CGAS

Developmentally Disabled Children's Global Assessment Scale (DD-CGAS)

Instructions for Raters

Areas that need to be considered in ratings include:

- Overall functioning in major adaptive domains:
 - Self care: eating, dressing, sleeping
 - Communication
 - Social behavior
 - Academic performance and setting
- Consistency or inconsistency of functioning across settings: home, school, community
 - Level of environmental adaptation needed
 - Level of supervision needed

1. Use the table below to organize your judgment of impairment across the four domains of function.

2. Choose the header/category that best describes general functioning (ex: *"moderate impairment in functioning in most areas"*). The descriptor should be a good description of the general functioning of the child, regardless of whether the source of impairment is cognitive, behavioral or other. You are comparing the description of adaptive functioning to what would be expected of a typically developing child, regardless of whether the impairment is due to developmental disability, behavioral disturbance, environmental factors, or other. Be wary of placing too much emphasis on standard scores; variability in functioning may get "averaged" out in the standard score. Place more emphasis on descriptions of functioning in the vignette.

3. Check details of that category to confirm that this is a general description, but note that most children will not fit perfectly into any particular category, you are looking for the "best fit".

4. When you think you have found the best fit, look at the two adjacent categories, to see if the child has some characteristics that fit into the next higher or lower category. This will help you adjust your score. For example, if the child fits best into *"60-51 Moderate impairment in functioning in most areas"* but has some similarity to 41-50, you would score in the lower half of the range (51-55). Conversely, if the child fits best in 60-51 but has some strengths that are consistent with the next higher category, you would score in the top half of the category (55-60).

		Level of Impairment				
Domain		None	Slight	Moderate	Severe	Extreme
	Self Care					
	Communication					
	Social Behavior					
	School/Academic					

Suggested reference: Wagner A, Lecavalier L, Arnold LE, Aman G, Scahill L, Stigler KA, Johnson C, McDougle CJ, Vitiello B. Developmental Disabilities Modification of the Children's Global Assessment Scale (DD-CGAS). *Biological Psychiatry*, vol XX, yyy-yyy.

Note. Readers are permitted to make free copies, as required. Electronic copies of the DD-CGAS can be obtained by writing to the authors.

Figure 2.
Scoring instructions

Table 1**Subject characteristics**

	Pilot (n = 17)	Samples RCT (n = 66)	Combined (n = 83)
Demographic			
Age (Years; Mean, SD)	8.1 (2.7)	7.5 (2.5)	7.6 (2.5)
Gender (% of boys)	82.3	77.2	78.3
Diagnosis (%)			
Autistic Disorder	11 (65)	45 (68)	56 (67)
Asperger Syndrome	2 (12)	4 (6)	6 (7)
PDD-NOS	4 (24)	17 (26)	21 (25)
Verbal children (%)	12 (71)	49(74)	61 (73)
Vineland – Composite Mean (SD)	41.5 (11.9) ^b	50.2 (16.5)	48.5 (16.0) ^a
Stanford Binet-5	-	70.0 (22.9) ^c	-
Leiter International-R	-	73.5 (22.9) ^d	-
ABC – Irritability (Mean, SD)	24.3 (9.2)	29.5 (6.3)	28.4 (7.2)
CY-BOCS – Total (Mean, SD)	13.8 (3.1)	15.5 (2.5)	15.2 (2.7)
Autism Symptomatology			
ADI-R Social	22.1 (4.6)	21.1 (6.2)	21.3 (5.9)
ADI-R Communication-V	16.2 (3.5) ^f	14.3 (5.1) ^g	14.7 (4.8) ^e
ADI-R Communication- NV	11.8 (4.4) ⁱ	11.5 (2.9) ^j	11.6 (3.2) ^h
ADI-R Repetitive	6.6 (2.9)	5.8 (2.6)	6.0 (2.6)
DD-CGAS (Mean, SD)	46.3 (11.0)	48.5 (11.0)	48.1 (10.9)

^a
n = 82;^b
n = 16;^c
n = 46;^d
n = 58^e
n = 61;^f
n = 12;^g
n = 49;^h
n = 22;ⁱ
n = 5;^j
n = 17.

Note. RCT=randomized controlled trial; PDD-NOS=pervasive developmental disorder, not otherwise specified; ABC=Aberrant Behavior Checklist; CY-BOCS=Children's Yale-Brown Obsessive Compulsive Scale; ADI-R=Autism Diagnostic Interview-Revised; DD-CGAS=Developmental Disabilities-Children's Global Assessment Scale

Table 2

Pearson correlation coefficients between DD-CGAS scores and other measures of symptoms and functioning

Measures	r	p	n
Adaptive behavior			
VABS-Composite ^b	.50	<.001	82
ABLLS ^a	.52	<.001	83
Intellectual Functioning			
Stanford-Binet 5 ^c	.47	.001	46
Leiter-R ^d	.49	<.001	58
Psychopathology			
ABC-Irritability ^a	-.30	.006	83
CYBOCS ^a	-.29	.008	83
HSQ ^a	-.26	.016	83
Autistic Symptomatology			
ADI-R Social ^a	-.30	.005	83
ADI-R Communication-Verbal ^e	-.09	.485	61
ADI-R Communication-Nonverbal ^f	-.45	.037	22
ADI-R Repetitive Behavior ^a	-.03	.797	83
CGI-Severity	-.48	<.001	83

Note. DD-CGAS=Developmental Disabilities-Children's Global Assessment Scale; VABS=Vineland Adaptive Behavior Scale; ABLLS=Assessment of Basic Language and Learning Skills; ABC= Aberrant Behavior Checklist; CY-BOCS=Children's Yale-Brown Obsessive Compulsive Scale; HSQ=Home Situations Questionnaire; ADI-R=Autism Diagnostic Interview-Revised