Sensitivity of Tests to Assess Improvement In ADHD Symptomatology

Natalie Grizenko, MD[1,2], Martin Lachance[1], Vincent Collard[1], Philippe Lageix, MD[1,2], Chantal Baron, MD[1,2], Leila Ben Amor, MD[1,2], Marina Ter Stepanian, BA[1], Valentin Mbekou, PhD[1], George Schwartz, M.Sc[1,2], Johanne Bellingham[2], Ridha Joober, MD[1,2]

ABSTRACT

Objective: To assess which measurements best predict improvement on ADHD symptomatology after medication is given. **Methods:** 147 children aged 6 to 12 years, diagnosed with ADHD, participated in a double-blind placebo controlled two-week crossover trial of methylphenidate. **Results:** There were statistically significant differences on all measures between placebo and medication. Effect size for the overall group was 0.33 (CGI-P), 0.80 (CGI-T), 1.33 (CGI), 0.56 (CPT), 0.82 (RASS). **Conclusions:** Acute behavioural response measures, where children are observed by clinicians (RASS and CGI), were overall more reliable than parent reports in detecting improvement on ADHD symptomatology. Teacher reports were also very important, especially in the 9 to 12 year old group.

Keywords: Attention Deficit Hyperactive Disorder, medication response, methylphenidate

RÉSUMÉ

Introduction: L'objectif de cette étude est de déterminer les méthodes les plus efficaces pour mesurer l'amélioration des symptômes du trouble déficitaire de l'attention avec hyperactivité (TDAH) après médication. **Méthodologie:** Cent quarante-sept enfants, âgés de 6 à 12 ans et souffrant du TDAH, ont participé à une étude de deux semaines à double insu contrôlée par placebo où chacun des groupes recevait, en alternance, la substance active ou le placebo. **Résultats:** On constate des différences significatives entre le placebo et la substance active. L'ampleur de l'effet sur la totalité du groupe est de 0,33 (CGI-P), 0,80 (CGI-T), 1,33 (CGI), 0,56 (CPT) et 0,82 (RASS). **Conclusions:** Les données obtenues à partir de réponses comportementales aiguës où les enfants ont été observés par des cliniciens (RASS et CGI) sont, dans l'ensemble, plus fiables que les observations des parents pour déceler une amélioration des symptômes du TDAH. Les évaluations des professeurs jouent également un rôle très important, notamment chez les enfants de 9 à 12 ans. **Mots-clés:** trouble déficitaire de l'attention avec hyperactivité, réponse à la médication, méthylphénidate.

INTRODUCTION

Attention deficit hyperactive disorder (ADHD) is very common amongst school-aged children, affecting 8.4% of boys and 3.1% of girls (Breton et al., 1999). Children with ADHD may present with decreased attention span, impulsiveness, restlessness and emotional instability. If not treated, the symptoms have a serious impact on academic functioning (Merrell & Tymms, 2001).

Methylphenidate (MPH) is the most commonly used medication in ADHD; 70% of treated patients respond to the medication. Other medications have also been used to treat ADHD such as dexedrine, desipramine, buproprion and clonidine, but there are numerous side effects associated with these medications (NIH, 2000). For efficient treatment of children with ADHD it is therefore important to determine the effectiveness of methylphenidate in improving symptomatology before switching to other treatment modalities.

A full array of tests have been used to evaluate improvement of symptoms of ADHD with MPH. In the ecological milieu the Conners' Global Index parent and teacher forms (CGI-P and CGI-T; Conners et al., 1998) have been utilized. Acute behavioural response can be assessed using the Clinical Global Impression Scales (CGI; Rapoport et al., 1985), the Continuous Performance Task (CPT; Conners, 1995) and the Restricted Academic Situation Scales (RASS; Barkley, 1990). Some of the limitations of these scales include: their subjective nature, for example, in the CGI-P each parent has his own interpretation of scale items. Secondly, scales commonly detect only certain factors. ADHD is a complex disorder with at least two major clinical dimensions: hyperactivity/impulsivity and inattention. Different tests might be more efficient with different age groups since older children present with more inattentive symptoms while younger children present with more symptoms of motor hyperactivity (Lahey et al., 1994). A global and objective test assessing all aspects of the disorder is lacking. Therefore, it is of great importance to know which test or combination of tests best detect how specific groups of children respond to MPH. Thirdly, most studies in the past have been directed toward finding the best scales to make an objective diagnosis of ADHD but even if a scale is very reliable in correlating with a correct diagnosis of ADHD, one cannot infer that the same scales would detect improvement with treatment.

A study by Fisher and Newby (1998), with a similar double blind placebo design to ours, showed that RASS was able to detect behavioural changes induced by medication, but lower

[1]Department of Psychiatry, McGill University; [2]Division of Child and Adolescent Psychiatry, Douglas Hospital
Corresponding author: grinat@douglas.mcgill.ca

and higher doses of MPH could not be differentiated only on the basis of RASS scores. A limitation of this study is that it did not compare the RASS with any other scales.

In a review published by Rapport, Chung et al., (2000), a small subset of acute behavioural response measures, including the CPT and the RASS, used to monitor treatment response in children appeared not to mirror the overall behavior and academic functioning in classroom setting. Moreover, all of the clinic-based measures failed to differentiate between various dosages.

In summary, no placebo-controlled trial has so far compared the ability of both ecological and acute behavioural response measures to detect the effects of a stimulant drug. The purpose of this study is to examine which tests or combination of tests are better able to predict improvement in symptomatology of ADHD in different age groups after treatment with methylphenidate.

METHODS

The present study is a double-blind placebo-controlled crossover trial. All baseline assessments of the children examining their degree of behavioural problems, IQ, academic performance, and severity of illness were carried out in the week prior to the trial. Children were assigned randomly to receive either one week of the active medication followed by one week of placebo or vice versa. Children randomly received either a placebo or 0.5 mg/kg/day of methylphenidate in a divided dose, which is the best-documented dose prescribed in clinical and research settings (Sprafkin & Dadow, 1996). Methylphenidate takes effect shortly after its administration. However, to make sure that slower treatment responses would not be missed, the children received medication for seven consecutive days. Both drug and placebo were prepared by a pharmacist in identical coloured gelatine capsules.

During the week of baseline evaluations, the children were assessed using the Conners' Continuous Performance Task (CPT) and the 10-item Conners' Global Index parent and teacher form (CGI-P and CGI-T). The next two week medication was administered daily in the morning and at noon. On the third day of each treatment week (testing day), before taking the morning medication, the child was evaluated using the CPT, the CGI scale and the RASS and then re-evaluated on the same measures one hour later to assess the acute effect of MPH versus placebo. CGI-P and CGI-T were completed at the end of each week reflecting child's overall performance during the preceding week.

Participants

The sample consisted of 126 boys and 21 girls, aged between 6 and 12 years. The average age of the group was 8.98 years (SD = 1.77). All of them met the DSM-IV diagnosis of ADHD. In all cases, a best estimate diagnosis of ADHD was made by a panel of child psychiatrists based on the Diagnostic Interview Schedule for Children-IV (DISC-IV; Shaffer et al., 2000) and on a clinical interview including school and parental reports. ADHD subtypes were diagnosed using the DISC-IV. Children currently taking medication other than MPH were excluded. Exclusion criteria also included having a history of Autism, Tourette syndrome, Pervasive Developmental Disorder, Psychosis and an IQ less than 70. The children selected for the study had a mean IQ of 96.5 (SD = 13.6) as measured by the Wechsler Intelligence Scale for Children-III (WISC-III; Wechsler, 1991).

Setting

All children meeting inclusion criteria were recruited sequentially from the Disruptive Behaviour Disorders Program and from the general outpatient services at the Douglas Hospital in Montreal, a psychiatric university teaching hospital. Ninety five percent of the eligible patients agreed to participate in the study. All children were enrolled in the study immediately after assessment and hence did not receive any significant amount of psychosocial intervention prior to the medication trial.

Measures

A) Ecological measures:

1. 10-item Conners' Global Index parent and teacher forms (CGI-P and CGI-T; Conners et al., 1998):

These scales assess children's restlessness, impulsiveness and emotional stability. Teachers and parents determine the frequency of occurrence of ten types of behaviours, which could have been seen in the preceding week, e.g., temper outbursts, fidgeting, etc. The scoring method is sex and age specific. Internal reliability coefficient for the CGI-P and the CGI-T are both 0.94, showing high consistency of all items. Test-retest reliability coefficients, over a 6 to 8 week interval, are 0.72 for the Parent form and 0.80 for the teacher form.

B) Acute behavioural response measures

1. Clinical Global Impression Scale (CGI; Rapoport et al., 1985):

The severity of illness and global improvement of the children was assessed by a research psychologist while they were undergoing the testing. The CGI ranges from 0 to 7, with 7 among the most extremely symptomatic patients. There was good inter-rater reliability.

2. Conners' Continuous Performance Task (CPT; Conners, 1995):

This computerized test measures response inhibition and impulse control as well as sustained attention. Letters are displayed at different intervals on the screen during 14 minutes. The child is instructed to press the space bar only when a letter other than X is shown. The CPT overall index is a weighted measure of different parameters including omission errors, commission errors, and time of response.

The Conner's CPT has been shown to provide a good means for monitoring the effectiveness of treatment (Conners, 1995). As described in its manual, there is a clear positive linear effect of dose of MPH on reaction time ($F = 9.81$, $p<.01$).

3. Restricted Academic Situation Scale (RASS; Barkley, 1990):

This task provides information about the frequency and severity of ADHD symptoms during performance of independent academic work. The child is left alone in a room with a set of math problems adapted to his academic level and told to do as many as he/she can in 15 minutes.

The child's behaviour is scored by a researcher through a one-way mirror over consecutive 15-second intervals on five behavioural categories: off-task, fidgeting, vocalizing, playing with objects and out of seat. The RASS has been shown to significantly discriminate children with ADHD from normal children (Milich, Loney & Landau, 1982). Previous research has also shown that the RASS is sensitive to improvement in scores with dosages as low as 0.2 mg/kg of methylphenidate. There was good inter-rater reliability.

Statistical Analysis

Children were divided into two age groups, 6 to 8 years old (n = 78) and 9 to 12 years old (n = 69).

Two-tailed paired t-tests were used to determine statistically significant change between placebo and active weeks. CPT and RASS change scores (second - first assessment) during a given testing day, representing acute effects of MPH or placebo, were analysed.

The effect size of a test represents the difference between the means of the patients on placebo and active medication over the SD of placebo. Cohen (1988) has described an effect size of 0.2 as being small, 0.5 as medium and 0.8 as large.

Placebo effect on the CGI-P and the CGI-T were determined from changes in scores from baseline to placebo only for those children who received the placebo in the first week of trial (n = 75).

Practice effects of acute behavioural response tests (RASS and CPT) were determined by examining the acute effect on the third day of the placebo week.

RESULTS

Demographic characteristics

Forty-six percent of the children came from a family with an income of less than $20,000 ($CAN) /year (low income), 26% of families earned between $20,000 and $40,000 /year (lower-middle income) and 26% earned more than $40,000 /year.

Fathers had an average education level of 11.5 years (SD= 3.3), mothers had an average of 12.2 years (SD= 2.9). Sixty-three percent of the children were diagnosed as having the combined form of ADHD, 26% the predominantly inattentive type and 11% the predominantly hyperactive type.

The children of the inattentive type had a mean age of 9.3 years old (SD = 1.9). They were on average older than the other two types. The mean age of the children diagnosed with the combined type or with the hyperactive type had a mean age of 9.0 years (SD = 1.7) and of 8.3 years (SD = 1.7) respectively.

Sixty percent of the children had received some kind of medication treatment prior to the study. All medication was discontinued for a period of 2 weeks prior to the start of the trial. None of the children dropped out of the study.

The DISC-IV (Shaffer et al., 2000) was used to evaluate the presence or absence of comorbidity. In our sample, 38.3% of children had an oppositional defiant disorder (ODD), 5.1% had a mild to severe conduct disorder (CD), 32.7% of children had both ODD and CD, 4.3% of the sample was assessed as having a generalized anxiety disorder and 5.2% as having a major depressive disorder.

Sensitivity of tests to detect improvement

There is a significant difference on all outcome measures between placebo and medication scores (Table 1). Furthermore, all five measures showed that both age groups, that is, the 6 to 8 year old group and the 9 to 12 year old group, did significantly better on medication than on placebo.

The effect size for the two age groups is presented in Table 2. Overall, the CGI-P has the smallest effect size, while the biggest effect size is on the CGI.

On average, all children, whether they received placebo or MPH in the first week of the study, improved their CGI-P score significantly from baseline to week 1 (Table 3). On the other hand, only the group receiving MPH in the first week of the trial improved their CGI-T score significantly from baseline to week 1.

No practice effects were found on acute behavioural response measures. In fact, the mean RASS scores were 11.3% worse during the second compared to the first testing session on the third day on placebo; the CPT the scores were 12.3% worse.

Since girls represented only 17% of the sample, analyses were repeated only with boys. The results were very similar and are not presented.

DISCUSSION

The study is unique in that it examines how different measures detect improvement in symptoms of children with ADHD when administered MPH. All outcome measures used showed a clinically significant improvement on MPH as compared to placebo.

In both the younger and older age group the CGI was the best measures to detect symptoms improvement on methylphenidate. Both symptoms of motor hyperactivity, which were more frequently found in the younger group, and inattention, which were more frequently found in the older group were captured by the CGI. The RASS was also effective in detecting changes in hyperactivity level (i.e. fidgeting) and inattention (i.e off task behaviour) in both age groups. Hence with two simple measures of acute behavioural response one can detect improvement on methylphenidate very accurately. This is clinically extremely relevant because it shows that a clinician can assess accurately response to medication by comparing two short visits - one on placebo and one on methylphenidate.

The study further demonstrates the CGI-T is much more accurate then the CGI-P in detecting symptoms improvement on medication. This can be explained by the very high placebo response when the parents complete the CGI-P. There was no specific placebo response in the teacher reports. Secondly, CGI-T had the second highest effect size in the 9 to 12 year old, indicating that teachers are quite good at detecting improvement in the older group who tend to present with more inattentive symptoms than the younger group who tend to be more hyperactive, a symptom that is more easily detected by observers on the RASS.

Finally the CPT showed only a moderate effect size in both age groups. This can be explained in that may of the children had comorbid oppositional defiant disorder and showed a reluctance to following the instructions in CPT. Furthermore the children frequently stated that they were bored by the CPT and answered haphazardly.

Strengths and Limitations

One of the main strengths of our project is the thorough and rigorous diagnostic evaluation carried out. The possibility of having false-ADHD children amongst our subjects is therefore very low. Moreover, because the children were referred from both the general out-patient program and the disruptive behaviour disorders program for children with severe difficulties, we were able to assess the ability of the tests to detect improvement in children across the full spectrum of ADHD symptomatology, from the mildly to the very severely ill.

Our analysis has been limited by the fact that our sample included too few girls to separate the data by sex or by different subtypes of ADHD, i.e. inattentive versus hyperactive-impulsive type.

Also, our study design only included one dose of methylphenidate. Therefore, the sensitivity of the various tests in relation to multi-dosing treatment remains a factor that should be further examined.

Clinical implications

In determining the responder status of a child with ADHD, tests may contradict one another. It is therefore of primary importance to know which tests should be given more weight in such an evaluation.

Since MPH is considered the treatment of choice for ADHD, it is crucial to have the best tool to detect improvement in responders to MPH. This would allow for a 1) proper titration (careful initial titration has been shown by the Multimodal Study of ADHD of the National Institute of Mental Health to be essential for an optimal pharmacological treatment of ADHD; (Vitielo et al., 2001), 2) help clinicians choose alternative methods of treatment such as dextroamphetamine, if response to methylphenidate is suboptimal, 3) provide parents with an objective assessment of the degree of clinical improvement in their children with medication and thus improve compliance.

ACKNOWLEDGEMENTS

The authors would like to thank the FRSQ and MRC for their financial support.

REFERENCES

Barkley, R., A. (1990). Attention Deficit Hyperactivity Disorder: *A Handbook for Diagnosis and Treatment*. New York: The Guilford Press.

Breton, J.J., Bergeron, L., Valla JP., Berthiaume C., Gaudet N., Lambert J., St-Georges M., Houde L. & Lepine S. (1999). Quebec Child Mental Health Survey: Prevalence of DSM-III-R Mental Health Disorders. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 40, 375-384.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences. New York: Academic Press.

Conners,C., K. (1995). Conners' Continuous Performance Test Computer Program. Toronto : MHS.

Conners, C., K., Sitarenios, G., Parker, J.D., & Epsteins, J.N. (1998). Revision and restandardization of the Conners Teacher Rating Scale: factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology*, 26, 279-291.

Conners, C., K., Sitarenios, G., Parker, J.D., & Epsteins, J.N. (1998). The revised Conners' Parent Rating Scale : factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology, 26*, 257-268.

Fisher, M., & Newby, R.B. (1998). Use of the Restricted Academic Task in ADHD Dose-Response Relationships. *Journal of Learning Disabilities, 31*, 608-612.

Lahey, B.B., Applegate, B., McBurnett, K., Biederman, J., Greenhill, L., & Hynd, G.W. (1994). DSM-IV field trials for attention deficit hyperactivity disorder in children and adolescents. *American Journal of Psychiatry*, 151, 1673–85.

Merrell, C., & Tymms, P.B. (2001). Inattention, hyperactivity and impulsiveness: their impact on academic achievement and progress. *British Journal of Educational Psychology, 71*, 43-56.

Milich, R., Loney, J., & Landau, S. (1982). The Independent dimensions of hyperactivity and aggression : A validation with playroom observation data. *Journal of Abnormal Psychology, 91*, 183-189.

National Institutes of Health. (2000). National Institutes of Health Consensus Development Conference Statement : Diagnosis and Treatment of Attention-Deficit/Hyperactivity Disorder (ADHD). *Journal of American Academy of Child and Adolescent Psychiatry, 39*, 182-193.

Rapoport, J. (1985). Rating scales and assessment for use in paediatric psychopharmacology research : Clinical Global Impression. *Psychopharmacology Bulletin, 21*, 839-841.

Rapport, M.D., Chung, K.M., Shore, G., Denney, C.B., & Isaacs, P. (2000). Upgrading and technology of assessment and Diagnosis: Laboratory and Clinic-Based Assessment of Children With ADHD. *Journal of Clinical Child Psychology, 29*, 555-568.

Shaffer, D., Fisher, P., Lucas, C.P., Dulcan, M.K., & Schwab-Stone, M.E. (2000). NIMH diagnostic interview schedule for children version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 28-38.

Sprafkin, J., & Dadow, K.D. (1996). Double-blind versus open evaluations of stimulant drug response in children with attention-deficit hyperactivity disorder. *Journal of Child and Adolescent Psychopharmacology, 6*, 215-228.

Vitiello, B., Severe, J.B., Greenhill, L.L., Arnold, L.E., Abikoff, H.B., Bukstein, O.G., Elliott, G.R., Hechtman, L., Jensen, P.S., Hinshaw, S.P., March, J.S., Newcorn, J.H., Swanson, J.M., & Cantwell, D.P. (2001). Methylphenidate dosage for children with ADHD over time under controlled conditions: lessons from the MTA. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 188-196.

Wechsler, D. (1991). Wechsler Intelligence Scale for Children - Third Edition: Manual. San Antonio, TX. Psychological Corporation.

TABLE 1
Mean scores obtained on the five measurements

Test	Mean	(SD)	N	t	Sig.
CGI-P(P)	63.59	(13.98)			
			133	3.22	.002
CGI-P(A)	59.03	(12.79)			
CGI-T(P)	66.85	(13.18)			
			117	7.97	<.001
CGI-T(A)	56.44	(11.69)			
CGI-(P)	4.54	(0.87)			
			93	8.26	<.001
CGI-(A)	3.39	(1.18)			
CPTd-(P)	2.07	(6.64)			
			133	5.34	<.001
CPTd-(A)	-2.41	(6.95)			
RASSd-(P)	6.11	(21.15)			
			146	8.48	<.001
RASSd-(A)	-17.09	(28.73)			

Note: CGI-P = 10-item Conners' Global Index parent form; CGI-T = 10-item Conners' Global Index teacher form; CGI = Clinical Global Impression global improvement; CPTd = Continuous Performance Test change score (time 2 - time 1); RASSd = Restricted Academic Situation Scale change score; P = placebo week; A = active week.

TABLE 2
Effect sizes of active medication

Test	**Type of measure**	**Whole Group (N = 147)**	**6 to 8 years old (N = 78)**	**9 to 12 years old (N = 69)**
CGI-P	Ecological	0.33	0.25	0.41
CGI-T	Ecological	0.80	0.81	0.78
CGI	Acute Behavioural Response	1.33	1.44	1.14
CPT	Acute Behavioural Response	0.56	0.56	0.57
RASS	Acute Behavioural Response	0.82	0.96	0.71

TABLE 3
Improvement from baseline to the first week of the trial

Week 1 assignment	Test	Mean Baseline score	Mean Week 1 score	Improvement (Relative%)	Sig.
Placebo	CGI-P	75,0	62,6	16,5	<.001
(N=75)	CGI-T	69,2	67,0	3,2	0.09
Active medication	CGI-P	75,2	57,9	23,0	<.001
(N=72)	CGI-T	71,0	57,9	18,5	<.001