Relations Between Multi-Informant Assessments of ADHD Symptoms, DAT1, and DRD4

Ian R. Gizer, Irwin D. Waldman, and Ann Abramowitz
Emory University

Yu Feng, Karen G. Wigg, and Virginia L. Misener
Toronto Western Research Institute

Cathy L. Barr
Toronto Western Research Institute and Hospital for Sick Children

David C. Rowe
University of Arizona

Researchers conducting candidate gene studies of attention-deficit/hyperactivity disorder (ADHD) typically obtain symptom ratings from multiple informants (i.e., mothers, fathers, and teachers) and use a psychologist’s best estimate or a simple algorithm, such as taking the highest symptom ratings across informants, to construct diagnostic phenotypes for estimating association. Nonetheless, these methods have never been empirically validated in the context of a molecular genetic study. In the current study, the authors systematically evaluated several methods of operationalizing phenotypes and the resulting evidence for association between ADHD and the candidate genes: dopamine transporter gene (DAT1) and dopamine D4 receptor gene (DRD4). Use of symptom scores as continuous scales in regression analysis suggested that the combination of mother and teacher ratings yielded the strongest evidence for association between hyperactive–impulsive ADHD symptoms and DAT1 and between inattentive ADHD symptoms and DRD4. Teacher ratings alone were sufficient for evaluating the association between inattentive symptoms and DAT1. Further, this regression-based method consistently yielded stronger evidence for association among ADHD symptoms, DAT1, and DRD4 than did three simple algorithms (i.e., the and, or, and averaging rules). The implications of these results for future molecular genetic studies of ADHD are discussed.

Keywords: ADHD, DAT1, DRD4, multiple informants

Attention-deficit/hyperactivity disorder (ADHD) is estimated to occur in approximately 3%–7% of children, making it one of the most prevalent childhood psychiatric disorders (American Psychiatric Association, 2000). As described in the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM–IV; American Psychiatric Association, 2000), ADHD consists of two distinct but correlated inattentive and hyperactive–impulsive symptom dimensions. Quantitative behavior genetic studies (i.e., twin and adoption studies) have suggested substantial genetic influences are involved in the etiology of ADHD with heritabilities ranging from 60%–90% (Waldman & Rhee, 2002). As a result, researchers have begun to search for susceptibility genes for this disorder with some success (see Faraone et al., 2005, and Waldman & Gizer, 2006 for reviews). Nonetheless, all initial reports claiming a relation between a specific candidate gene and ADHD have been followed by a mix of successful and failed replication attempts. It is important to note that many of these studies have identified such susceptibility genes by testing for an association between specific candidate genes and ADHD defined as a discrete diagnostic category. Such an approach is warranted, given that twin studies suggest substantial genetic influences on ADHD operationalized as a diagnostic category (Thapar, Holmes, Poulton, & Harrington, 1999), but similar studies also suggest that the continuous, inattentive, and hyperactive–impulsive symptom dimensions that underlie ADHD show similarly strong genetic influences (e.g., Sherman, Iacono, & McGue, 1997). Thus, researchers have begun to explore whether the susceptibility genes that have been identified for the diagnosis of ADHD might be better conceptualized as quantitative trait loci (QTLs) that underlie the hyperactive–impulsive and/or inattentive symptom dimensions (see Asherson & Image Consortium, 2004 for a review). Given the statistical power that is retained when one does not artificially dichotomize a continuum, the QTL approach may provide a more powerful method for identifying susceptibility genes (Cohen, 1983).

Although the QTL approach may present specific advantages for examining the association of candidate genes with ADHD, studies of this type present unique difficulties regarding the as-
essment of hyperactive–impulsive and inattentive symptoms. The diagnostic criteria for ADHD require that presenting symptoms be a pervasive aspect of the child’s behavior and, therefore, observable in at least two settings (e.g., home and school) (American Psychiatric Association, 2000). Thus, researchers often assess ADHD symptoms by collecting data from multiple informants—most often the child’s parents and teachers—in order to establish pervasiveness. Nonetheless, researchers have shown via meta-analysis that the correlations between informants’ symptom ratings are low (e.g., parent–teacher, r = .27; Achenbach, McConaughy, & Howell, 1987). Studies conceptualizing ADHD as a discrete diagnostic category often account for such discrepancies by use of best-estimate procedures for combining assessment data from multiple informants in a manner that mirrors the decision process clinicians use when determining clinical diagnoses. For studies with a QTL approach, however, a similar method does not exist and may not be preferable. As a result, specific methods have been used in studies examining the association between a candidate gene and a continuously distributed trait, such as taking the highest score among informants (e.g., Rowe et al., 1998) or taking an average score across informants (e.g., Mill et al., 2005), but these methods have yet to be systematically evaluated.

Investigations of methods for use of multi-informant data in molecular genetic studies are informed by two important lines of research. First, many studies have demonstrated that parent and teacher reports provide unique and valid diagnostic information (see Achenbach, McConaughy, & Howell, 1987 and Meyer et al., 2001 for reviews), and further studies have demonstrated the incremental usefulness of having multiple informants to diagnose ADHD (e.g., Power et al., 1998) as well as diagnosing ADHD subtypes in research settings (e.g., Gadow et al., 2000; Weiler, Bellinger, Marmor, Rancier, & Waber, 1999). Additional research has suggested that combining assessment data from multiple informants into a single clinical picture can obscure important clinical data. For example, Gadow et al. (2004) demonstrated that parent and teacher ratings of ADHD differ in terms of their sensitivity to ADHD symptoms as well as their relations with measures of social, academic, and cognitive functioning. This led the authors to conclude researchers studying childhood psychiatric disorders should evaluate informant-specific diagnoses as well as diagnoses derived from the combination of informant reports.

Second, and of direct relevance to genetic association studies, twin studies of ADHD suggest that parent and teacher ratings of ADHD symptoms show genetic influences that are unique to each informant, in addition to those that are common across informants (Martin, Scourfield, & McGuffin, 2002; Nadder, Silberg, Rutter, Maes, & Eaves, 2001). For example, one study reported that although approximately 31% of the variability in parent and teacher ratings of ADHD was due to a common genetic etiology, as much as 40% of the variability in parent ratings and 50% of the variability in teacher ratings were due to genetic influences unique to that informant’s ratings (Martin, Scourfield, & McGuffin, 2002). This suggests that although parent and teacher ratings are to some extent measuring a common trait influenced by a shared set of genes, each informant’s ratings are also measuring distinct aspects of ADHD symptomatology influenced by distinct genes.

Such findings have obvious implications for molecular genetic studies of ADHD. If different informants’ symptom ratings reflect specific as well as common genetic influences, studies that test for association with phenotypes based on a single informant could result in negative findings or in failures to replicate previous findings if they do not have the same informant as the initial study. Further, studies that combine data across informants without careful consideration of the methods used might be losing valid diagnostic data, which could also lead to negative findings. Thus, testing for association for each informant separately and in combination should provide a clearer understanding of the relation between a candidate gene and a disorder. In the current study, ADHD symptom scale scores were used to evaluate the unique and incremental effects of mother, father, and teacher ratings of ADHD, as well as the interactions between informant ratings, in their associations with two candidate genes that have shown replicable evidence of association with ADHD, the dopamine transporter gene (DAT1) and the dopamine D4 receptor gene (DRD4).

DAT1 was first suggested as a candidate gene for ADHD given that stimulant medications prescribed to treat ADHD symptoms, such as methylphenidate, appear to bind to the dopamine transporter and block the reuptake of dopamine from the synapse thereby increasing the amount of available dopamine (Ritz, Lamb, Goldberg, & Kuchar, 1987; Volkow et al., 1995). Cook et al. (1995) conducted the first study to test for association between DAT1 and ADHD, and they reported evidence of association between a 40 base pair (bp) variable number of tandem repeats (VNTR) sequence in the 3′ untranslated region (UTR) of DAT1 and ADHD. More specifically, they found that the 10-repeat allele was preferentially transmitted to children with ADHD. This relation has been replicated several times (e.g., Curran et al., 2001; Gill, Daly, Heron, Hawi, & Fitzgerald, 1997; Waldman et al., 1998), but there have also been several failures to replicate (e.g., Holmes et al., 2000; Todd et al., 2001). Nonetheless, a recent meta-analysis of published and unpublished studies suggested a significant association between ADHD and the 10-repeat allele of DAT1 (Faraone and Kahn, 2006). These findings were extended by a study conducted in our own lab, suggesting that the severity of hyperactive–impulsive symptoms, but not inattentive symptoms, increased as the number of 10-repeat alleles (i.e., 0, 1, or 2 alleles) increased and, further, that DAT1 was associated with the combined ADHD subtype but not the inattentive ADHD subtype (Waldman et al., 1998).

Psychiatric genetic studies have also suggested an association between ADHD and DRD4. Initial interest in DRD4 was sparked by association studies linking the gene to the personality trait of novelty seeking (Benjamin et al., 1996; Ebstein et al., 1996). LaHoste et al. (1996) conducted the first test for association between ADHD and a 48-bp VNTR in exon 3 of DRD4. They reported an association between the seven-repeat allele of this polymorphism and ADHD. As with DAT1, this initial finding has been replicated several times (e.g., Barr et al., 2000; Faraone et al., 1999; Rowe et al., 1998), though there have also been several failures to replicate (e.g., Castellanos et al., 1998; Hawi et al., 2000). Faraone, Doyle, Mick, and Biederman (2001) conducted a meta-analysis of these and additional studies and found significant evidence suggesting that the seven-repeat allele of DRD4 is associated with increased risk for ADHD, which has remained a significant result even as the number of studies examining this relation has continued to increase (Faraone et al., 2005). These findings have also been extended by studies suggesting that DRD4 is more strongly associated with the inattentive than with the
The purpose of the current study was to evaluate alternative methods of using multi-informant data within the context of a molecular genetic study, specifically to evaluate the extent to which these methods provide evidence for an association between ADHD, DAT1, and DRD4. Because the majority of studies that examined the relation between ADHD and these two genes relied on best-estimate or consensus diagnoses to classify their samples, the current study provided an opportunity to evaluate different methods for using multi-informant data and to investigate the nature of the relations between these genes and ADHD.

Methods for using multi-informant data were evaluated in two steps. First, informants’ ratings of ADHD symptoms were evaluated with a regression-based optimal informant method (Bird, Gould, & Staghezza, 1992). In this approach, each informant’s symptom ratings are evaluated independently and in combination within a regression framework that includes tests for main effects and interactions. Entering these terms hierarchically into the regression model allows for the estimation of the incremental contribution of each informant’s symptom ratings and the interactions between informants’ ratings. Thus, the optimal informant or set of informants can be determined by testing whether each informant’s ratings or interaction of ratings yields a significant increase in explained variance when predicting the criterion variable (Bird et al., 1992). In the context of a molecular genetic study, mother, father, and teacher symptom ratings would be evaluated in terms of their association with a specific genetic marker. Such analyses were conducted as a first step to determine which informants’ ratings or interaction of ratings yields a significant increase in explained variance when predicting the criterion variable (Bird et al., 1992).

In the second step of the analyses, the results obtained with the regression-based method were contrasted against the evidence for association yielded by three simpler and more frequently used algorithms. The first algorithm, the or rule, designates a symptom as present if any informant rates it as present, thus providing a liberal estimate of diagnosis (Cohen, Velez, Kohn, Schwab-Stone, & Johnson, 1987; Lahey et al., 2000). When applied to continuous symptom scales, the or rule translates into use of the highest rating of each symptom dimension given by any informant. The second algorithm, the and rule, designates a symptom as present only if all informants (e.g., the parent and the teacher) rate it as present (Offord et al., 1996), which yields a conservative estimate of diagnosis. With continuous symptom scales, taking the lowest score given for each symptom dimension by each of the informants translates into a similarly conservative symptom rating. The third simple algorithm takes the average of the ratings for all available informants. Rather than simply the highest or lowest rating, with this method one considers each informant’s ratings, and therefore, this method represents a compromise between the or and the and algorithms.

Method

Participants

Ninety-nine children were recruited through the Center for Learning and Attention Deficit Disorders (CLADD) at the Emory University School of Medicine in Atlanta, Georgia, a clinic specializing in the assessment and treatment of childhood externalizing disorders such as ADHD, oppositional defiant disorder (ODD) and conduct disorder (CD). All children included in the study were probands referred to the clinic for an assessment related to externalizing behavior problems. Children diagnosed with autism, traumatic brain injury, or neurological conditions (e.g., epilepsy) were excluded from the study, as were children with IQs < 75. Any other diagnosis assigned to a child remained confidential and did not influence their inclusion in the study, though a post hoc analysis indicated that 77 of the 99 children met full criteria for ADHD. More specifically, 37 children met criteria for the combined subtype, 39 met criteria for the predominantly inattentive subtype, and 1 met criteria for the predominantly hyperactive–impulsive subtype. Further, 45 children met diagnostic criteria for a diagnosis of ODD, and 10 children met diagnostic criteria for CD. Diagnoses of internalizing disorders were not obtained for the sample.

The participants represent a subset of a larger sample, from which findings have been reported previously (e.g., Rowe et al., 1998; Waldman et al., 1998), as well as new participants in our ongoing study. It is important to note that the recruitment method did not change with the addition of the new participants, and further the participants did not significantly differ with respect to ADHD symptom levels as rated by mothers, fathers, or teachers (p values ranged from .293 to .686). They also did not differ with respect to age, gender, ethnic composition, or level of parental education (p values > .10). Participants ranged in age from 5 years to 17 years (M = 10.1, SD = 2.96) at the time of assessment. The sample consisted of 68 boys (68%) and 31 (32%) girls. The ethnic composition of the sample was 83% Caucasian, 8% African American, 8% Hispanic, and 1% Asian. The level of parental education was 2% with some high school, 10% who were high school graduates, 10% with some college but without a degree, 20% with a two-year college degree, 50% with a four-year college degree, and 8% with an advanced-level degree.

Assessment Procedures

Mother, father, and teacher ratings were obtained whenever possible for each child with the Emory Diagnostic Rating Scale (EDRS; Waldman et al., 1998). For those cases in which reports were obtained from multiple teachers, we used data from the teacher who spent the most time per week with the child. If there was no clear difference in the amount of time spent together, the homeroom teacher was selected. Because participants were not selected with respect to medication status, all informants were asked to rate the child’s behavior as though the child were not currently taking medication. Self-report data were not collected from the children.

The EDRS is a symptom checklist developed to assess symptoms of the major DSM-IV childhood psychiatric disorders, including disruptive disorders such as ADHD, ODD, and CD, and internalizing disorders such as depression and anxiety disorders. Parents and teachers rated symptoms on a 0–4 scale, with a score of 0 indicating that the symptom is not at all characteristic of the child and a score of 4 indicating that the symptom is very much characteristic of the child. A previous study showed that diagnoses derived by counting the number of symptoms present (i.e., ≥6
inattentive symptoms or ≥6 hyperactive–impulsive symptoms) yielded diagnostic rates in a control population similar to the prevalence rates described in the DSM–IV, suggesting the EDRS provides valid assessments of ADHD symptoms (Waldman et al., 1998). Internal consistencies of the hyperactive–impulsive scale, which contains the nine DSM–IV ADHD hyperactive–impulsive symptoms, and the inattentive symptom scale, which contains the nine DSM–IV ADHD inattentive symptoms, were independently evaluated for each informant to ensure that the scales possessed acceptable reliabilities. These values ranged from α = .83 to α = .96. The scores for each symptom were then summed to create hyperactive–impulsive and inattentive symptom dimensions (see Table 1 for means and standard deviations displayed by genotype).

Genotyping

The DNA collection and extraction procedures have been described in detail in previous publications (Rowe et al., 1998; Waldman et al., 1998). For the 40-bp VNTR in the 3' UTR of DAT1, samples were genotyped by polymerase chain reaction (PCR), with primers described by Vandenbergh et al. (1992), either according to the protocol described by Waldman et al. (1998) or, more recently, according to an alternate protocol. For the alternate protocol, PCR reactions were carried out in a 20 ul volume containing 60–100 ng DNA, 25 ng each primer, a set of Invitrogen PCR Enhancer System reagents consisting of 1X amplification buffer, 1X enhancer solution, 1.5 mM MgSO4 (Canadian Life Technologies, Burlington, Ontario, Canada), 0.2 mM dNTPs (0.05 mM each), and 0.5 units Taq polymerase, with thermocycling conditions as follows: (a) initial denaturation (4 min at 94°C), (b) 35 cycles of denaturation (40 sec at 94°C), annealing (40 sec at 68°C), and extension (30 sec at 72°C), and (c) final extension (10 min at 72°C). PCR products were resolved on 3% agarose gels and visualized by ethidium bromide staining. For the 48-bp VNTR in exon 3 of DRD4, samples were genotyped according to Lichter et al. (1993).

Analyses

To test for association and provide a direct test of the incremental contribution of each informant over and above every other informant, ordinal logistic regression analyses were conducted on each informant’s ratings uniquely and in combination with every other informant’s ratings. The number of high-risk alleles (i.e., number of 10-repeat alleles for DAT1 and number of 7-repeat alleles for DRD4) that a child possessed (i.e., 0, 1, or 2) served as the dependent variables in these analyses. Mother, father, and teacher ratings and interactions between these ratings served as independent variables. Participants’ ethnicities were included as covariates in these analyses to control for possible population stratification biases. Such biases occur when ethnic groups differ in both allele frequencies and rates (or symptom levels) of the disorder, which can produce spurious evidence for an association in the absence of any true relation between gene and disorder. Age and gender were also evaluated as potential covariates. The overall pattern of results were similar whether ethnicity, age, and gender were included as covariates with only slight changes in the effect sizes (data not shown). Thus, to simplify the presentation of findings, the results without covariates included are reported be-
low. Additionally, because directional a priori predictions were made regarding which allele would be associated with ADHD for each candidate gene, one-tailed \( p \) values were reported for the main effects. No strong predictions could be made in regard to the nature of the interactions between informant ratings in predicting genotype, and thus, two-tailed \( p \) values were reported for these results.

As individual predictors are entered hierarchically into an ordinal regression model, the Wald \( \chi^2 \) for the added predictor acts as a test of the incremental contribution of this predictor over and above previously entered predictors (Cohen, Cohen, West, & Aiken, 2003). Thus, to evaluate the incremental contribution of each rater and of each of the interaction terms, the significance of the Wald \( \chi^2 \) and the associated change in the Nagelkerke \( R^2 \) were evaluated for each step. This method determined whether the gain achieved by adding each successive predictor significantly improved the fit of the model or whether the results of the more parsimonious model were equivalent and, therefore, preferred. It should also be noted that the independent variables were mean-centered prior to calculation of the interaction term, to control for problems related to multicollinearity.

A key assumption underlying ordinal regression is that the slope of the regression line is approximately equal across each level of the dependent variable. To examine whether this assumption was met, the test of parallel lines was conducted for each analysis. When the assumption was violated, these results were reported, and binary logistic regression was used as an alternative. For DAT1, children with zero or one high-risk alleles were combined into a single group and contrasted against children with two high-risk alleles, given that very few children possessed zero copies of the high-risk allele. For DRD4, children with one or two high-risk alleles were combined into a single group and contrasted against children with zero high-risk alleles given that very few children possessed two copies of the high-risk allele.

In the current study, there were equal numbers of participating mothers and teachers (93 mothers and 93 teachers), but far fewer fathers (\( n = 67 \)). Thus, evaluating the incremental contribution of teacher ratings over and above mother ratings was accomplished by excluding a few participants due to missing data. Estimating the incremental contribution of father ratings was problematic, however, because almost one third of the sample (32 participants) would be excluded due to missing father data. Therefore, initial ordinal regression analyses were run with the full sample to estimate the effects of mother and teacher ratings uniquely and in combination, and a second set of analyses were conducted limiting the sample to participants with complete data from mothers, fathers, and teachers. The set of informants that yielded the best fitting model and demonstrated incremental contributions over a more parsimonious model was considered the optimal set of informants. As a final step, the best fitting model identified by regression analysis was evaluated against models with composite ratings produced by each of the simple algorithms (i.e., the or, and, and the averaging rules) by comparing the effect sizes yielded by each method.

### Results

Correlations between informant’s ratings were conducted to estimate the degree of overlap across informants (see Table 2). Mother and father ratings were highly correlated for both inattentive (\( r = .72 \)) and hyperactive–impulsive (\( r = .72 \)) symptoms. The degree of overlap was considerably lower between parent and teacher ratings for each dimension, ranging from \( r = .23 \) for father and teacher ratings of hyperactive–impulsive symptoms to \( r = .44 \) for mother and teacher ratings of hyperactive–impulsive symptoms.

As described in the Method section, estimating the incremental contribution of father ratings over and above mother or teacher ratings was problematic because father ratings could be collected from only two thirds of the sample. Further, analyses including and excluding father ratings yielded highly similar results for each candidate gene, with father ratings failing to provide any increase in explained variance over and above mother and teacher ratings (data not shown). Thus, only the results from analyses with mother and teacher ratings are reported.

**DAT1 and Hyperactive–Impulsive Symptoms**

To test for an association between hyperactive–impulsive ADHD symptoms and DAT1, mother and teacher ratings were evaluated uniquely and in combination with ordinal regression (see Table 3). Teacher ratings showed significant evidence of association with DAT1, Wald \( \chi^2(N = 93) = 3.38, p = .033 \), Nagelkerke \( R^2 = .045 \), whereas mother ratings did not, Wald \( \chi^2(N = 93) = 1.16, p = .140 \), Nagelkerke \( R^2 = .015 \). Further, teacher ratings continued to show incremental evidence for an association after

---

**Table 2**

<table>
<thead>
<tr>
<th>Informant/symptom scale</th>
<th>Mother inattentive sxs Correlation</th>
<th>n</th>
<th>Father inattentive sxs Correlation</th>
<th>n</th>
<th>Teacher inattentive sxs Correlation</th>
<th>n</th>
<th>Mother hyperactive–impulsive sxs Correlation</th>
<th>n</th>
<th>Father hyperactive–impulsive sxs Correlation</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Father inattentive sxs</td>
<td>.72b</td>
<td>69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher inattentive sxs</td>
<td>.25b</td>
<td>99</td>
<td>.30b</td>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother hyperactive–impulsive sxs</td>
<td>.41b</td>
<td>99</td>
<td>.25b</td>
<td>69</td>
<td>.26b</td>
<td>99</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father hyperactive–impulsive sxs</td>
<td>.41b</td>
<td>68</td>
<td>.41b</td>
<td>71</td>
<td>.24b</td>
<td>69</td>
<td>.72b</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher hyperactive–impulsive sxs</td>
<td>.04</td>
<td>99</td>
<td>-.06b</td>
<td>80</td>
<td>.51b</td>
<td>99</td>
<td>.44b</td>
<td>99</td>
<td>.23b</td>
<td>69</td>
</tr>
</tbody>
</table>

*Note.* Bold = significant correlation; sxs = symptoms.

*b* Correlation is significant at the .05 level (two tailed).  

*b* Correlation is significant at the .01 level (two tailed).
controlling for the presence of mother ratings, $\chi^2(N = 88) = 4.88$, $p = .014$, $\Delta R^2 = .062$, as indicated by the significant Wald $\chi^2$ test statistic. It was concluded that the best-fitting model, however, included the interaction between mother and teacher ratings. The improvement in fit over and above the main effects model approached significance, $\chi^2(N = 88) = 3.07$, $p = .080$, and resulted in a substantial increase in effect size ($\Delta R^2 = .042$).

A visual inspection of the described interaction indicated that teacher ratings of hyperactive–impulsive symptoms showed an increasing linear relation with the number of $DAT1$ high-risk alleles as mother ratings of these symptoms increased. This is indicated in Figure 1 by the increasing slopes of the regression lines, which were graphed with mother ratings set at $-1$ standard deviation, at the mean, and at $+1$ standard deviation. Because only 4 participants possessed zero copies of the $DAT1$ high-risk allele, this analysis was rerun excluding those participants to ensure they were not solely responsible for the interaction. The results remained highly similar after excluding these participants: interaction term, $\chi^2(N = 84) = 2.51$, $p = .056$.

This regression-based optimal informant model was then compared with the simple algorithms. As shown in Table 3, each of the simple algorithms yielded significant evidence for association between $DAT1$ and hyperactive–impulsive symptoms: or rule, Wald

Table 3
Association Analyses Between $DAT1$ and Mother and Teacher Ratings of Hyperactive-Impulsive Symptoms using Ordinal Regression

<table>
<thead>
<tr>
<th>Model</th>
<th>Overall model</th>
<th>Individual variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
<td>$\chi^2$</td>
</tr>
<tr>
<td>M only</td>
<td>93</td>
<td>1.165</td>
</tr>
<tr>
<td>T only</td>
<td>93</td>
<td>3.385</td>
</tr>
<tr>
<td>M &amp; T</td>
<td>88</td>
<td>5.678</td>
</tr>
<tr>
<td>M</td>
<td>88</td>
<td>.912</td>
</tr>
<tr>
<td>T</td>
<td>88</td>
<td>4.884</td>
</tr>
<tr>
<td>T &amp; M &amp; T</td>
<td>88</td>
<td>9.022</td>
</tr>
</tbody>
</table>

Simple algorithms

<table>
<thead>
<tr>
<th></th>
<th>Overall model</th>
<th>Individual variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>and rule</td>
<td>98</td>
<td>3.624</td>
</tr>
<tr>
<td>or rule</td>
<td>98</td>
<td>3.429</td>
</tr>
<tr>
<td>average method</td>
<td>98</td>
<td>3.580</td>
</tr>
</tbody>
</table>

Note. Boldface indicates significance at the .05 level. Italic represents a trend at the .10 level. The overall model $p$ is two tailed. The individual variables $p$ is one tailed, except where noted. For univariate regression analyses, individual variables $R^2$ represents the effect size for that model. For multiple regression analyses, $\Delta R^2$ represents the incremental contribution of the individual variable over and above the other variables in the model; OR = odds ratio; M = mother; T = teacher. $DAT1$ = dopamine transporter gene.

*a* Assumption of equal slope coefficients across categories was violated; statistics represent the results of logistic regression analysis.  
*b* Two tailed $p$.  

![Figure 1](image-url)

**Figure 1.** Graph of the interaction between mother and teacher ratings of hyperactive–impulsive (Hyper-Imp) symptoms in predicting the number of dopamine transporter gene ($DAT1$) high-risk alleles in which the linear relation between teacher ratings and $DAT1$ genotype is shown for three different values of mother ratings: the mean, $+1$ standard deviation, and $-1$ standard deviation.
RELATIONS BETWEEN MULTI-INFORMANT ASSESSMENTS

\[ \chi^2(N = 98) = 3.43, \ p = .032, \text{ Nagelkerke } R^2 = .042; \text{ averaging rule, Wald } \chi^2(N = 98) = 3.58, \ p = .029, \text{ Nagelkerke } R^2 = .044; \text{ and rule, Wald } \chi^2(N = 98) = 3.62, \ p = .028, \text{ Nagelkerke } R^2 = .051. \]

The test of parallel lines, which checks the assumption that the slope coefficients are equivalent across the levels of the dependent variable, was violated for the analysis of the and rule, \( \chi^2(1, N = 98) = 4.67, \ p = .031, \) and as a result, this algorithm was evaluated with a binary logistic regression in which participants with zero or one high-risk alleles were combined into a single group.

Despite the significant results produced by the simple algorithms, the regression-based optimal informant model that included mother and teacher ratings as well as their interaction yielded an effect size (i.e., Nagelkerke \( R^2 \)) that was \( \sim 2 \)–3 times as large as that yielded by the and rule and \( \sim 3 \) times as large as that yielded by the or and averaging rules. Thus, the optimal informant method appeared to yield stronger evidence for association between hyperactive–impulsive symptoms and DAT1 than did the simple algorithms.

**DAT1 and Inattentive Symptoms**

A similar set of analyses were conducted evaluating the relation between DAT1 and mother and teacher ratings of children’s inattentive symptoms (see Table 4). Teacher ratings of inattentive symptoms showed significant evidence for association with DAT1 when analyzed alone, Wald \( \chi^2(N = 93) = 7.26, \ p = .004, \) Nagelkerke \( R^2 = .100, \) but mother ratings did not, Wald \( \chi^2(N = 93) = .06, \ p = .400, \) Nagelkerke \( R^2 = .001. \) Further, teacher ratings continued to show incremental evidence for association after controlling for the presence of mother ratings, Wald \( \chi^2(N = 88) = 8.78, \ p = .002, \Delta R^2 = .131, \) and adding the interaction term into the model failed to yield a statistically significant increase in explained variance, Wald \( \chi^2(N = 88) = 1.24, \ p = .264, \Delta R^2 = .016. \) Thus, the model that included just teacher ratings of inattentive ADHD symptoms was found to be the best fitting.

This optimal informant model was then compared with each of the simple algorithms. As shown in Table 4, the optimal informant model including only teacher ratings yielded an effect size (Nagelkerke \( R^2 = .100) \) that was \( \sim 2 \)–3 times as large as that of any of the simple algorithms (Nagelkerke \( R^2 = .032–.046). \) Thus, the regression-based method appeared to yield stronger evidence for association between inattentive symptoms and DAT1 than did the simple algorithms.

**DRD4 and Hyperactive–Impulsive Symptoms**

The relations between mother and teacher ratings of ADHD hyperactive–impulsive symptoms and DRD4 were evaluated uniquely and in combination with ordinal regression (Table 5). Neither mother ratings nor teacher ratings of hyperactive–impulsive symptoms showed a relation with the number of DRD4 high-risk alleles, Wald \( \chi^2(N = 87) = 0.87, \ p = .176, \) Nagelkerke \( R^2 = .013, \) and Wald \( \chi^2(N = 89) = 1.46, \ p = .113, \) Nagelkerke \( R^2 = .023, \) respectively, nor did any of the additional models tested. Among the simple algorithms, only the or rule showed a trend suggesting an association between DRD4 and hyperactive–impulsive symptoms, Wald \( \chi^2(N = 92) = 2.65, \ p = .052, \) Nagelkerke \( R^2 = .043. \) It should be noted that because the test of parallel lines neared significance, \( \chi^2(1, N = 92) = 3.58, \ p = .059, \) for the analysis of the or rule, this algorithm was evaluated with a binary logistic regression in which participants with one or two high-risk alleles were combined into a single group. Thus, with the exception of the or rule, these results provide little evidence to suggest an association between hyperactive–impulsive symptoms and DRD4.

**DRD4 and Inattentive Symptoms**

Ordinal regression analyses were conducted to test the association between mother and teacher ratings of inattentive

---

**Table 4**

<table>
<thead>
<tr>
<th>Model</th>
<th>Overall model</th>
<th>Individual variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>( \chi^2 )</td>
</tr>
<tr>
<td>M only</td>
<td>93</td>
<td>1.055</td>
</tr>
<tr>
<td>T only</td>
<td>93</td>
<td>1.761</td>
</tr>
<tr>
<td>M &amp; T</td>
<td>88</td>
<td>10.021</td>
</tr>
<tr>
<td>M</td>
<td>11.309</td>
<td>3</td>
</tr>
<tr>
<td>T</td>
<td>M &amp; T</td>
<td>88</td>
</tr>
<tr>
<td>M</td>
<td>1.274</td>
<td>1.245</td>
</tr>
<tr>
<td>T</td>
<td>M &amp; T</td>
<td>1.502</td>
</tr>
<tr>
<td>M' T</td>
<td>1.508</td>
<td>2.610</td>
</tr>
<tr>
<td>Simple algorithms</td>
<td>98</td>
<td>1.629</td>
</tr>
</tbody>
</table>

Note. Boldface indicates significance at the .05 level. Italic represents a trend at the .10 level. The overall model \( p \) is two tailed. The individual variables \( p \) is one tailed, except where noted. For, univariate regression analyses, individual variables \( R^2 \) represents the effect size for that model. For multiple regression analyses, \( \Delta R^2 \) represents the incremental contribution of the individual variable over and above the other variables in the model; OR = odds ratio; M = mother; T = teacher; DAT1 = dopamine transporter gene.

a Two-tailed \( p \).
ADHD symptoms and DRD4 (Table 6). Teacher ratings were significantly associated with DRD4 when analyzed alone, Wald χ²(N = 90) = 5.28, p = .007, Nagelkerke R² = .088, whereas mother ratings were not, Wald χ²(N = 88) = 0.04, p = .578, Nagelkerke R² = .001. Further, teacher ratings continued to show incremental evidence for association after controlling for mother ratings, Wald χ²(N = 85) = 4.93, p = .013, ΔR² = .086. It was concluded that the best-fitting model, however, included the interaction between mother and teacher ratings. The improvement in fit over and above the main effects model was significant, Wald χ²(N = 85) = 4.34, p = .036, and resulted in a substantial increase in effect size (ΔR² = .082). A visual inspection of this interaction (see Figure 2) suggested that teachers’ ratings showed an increasing linear relation with the number of DRD4 high-risk alleles as mother ratings of these symptoms decreased. This is indicated by the decrease in slopes of the regression lines, which were graphed with mother ratings set at −1 standard deviation, at the mean, and at +1 standard deviation. Because only 2 participants possessed two copies of the DRD4 high-risk allele, this analysis was rerun excluding those participants to ensure they were not solely responsible for the significant interaction. The results remained highly similar after excluding these participants: interaction term, Wald χ²(N = 83) = 4.41, p = .036.

Table 5
Association Analyses Between DRD4 and Mother and Teacher Ratings of Hyperactive–Impulsive Symptoms Using Ordinal Regression

<table>
<thead>
<tr>
<th>Model</th>
<th>Overall model</th>
<th>Individual variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>χ²</td>
</tr>
<tr>
<td>M only</td>
<td>87</td>
<td>1.210</td>
</tr>
<tr>
<td>T only</td>
<td>89</td>
<td>1.241</td>
</tr>
<tr>
<td>M &amp; T</td>
<td>84</td>
<td>2.239</td>
</tr>
<tr>
<td>T</td>
<td>87</td>
<td>1.198</td>
</tr>
<tr>
<td>M &amp; T &amp; M’T</td>
<td>84</td>
<td>2.869</td>
</tr>
<tr>
<td>M</td>
<td>84</td>
<td>1.170</td>
</tr>
<tr>
<td>T</td>
<td>84</td>
<td>1.207</td>
</tr>
</tbody>
</table>

Note. Boldface indicates significance at the .05 level.Italic represents a trend at the .10 level. The overall model p is two tailed. The individual variables p is one tailed, except where noted. For univariate regression analyses, individual variables R² represents the effect size for that model. For multiple regression analyses, ΔR² represents the incremental contribution of the individual variable over and above the other variables in the model; OR = odds ratio; M = mother; T = teacher. DRD4 = dopamine D4 receptor gene.

a Two tailed p. b Assumption of equal slope coefficients across categories was violated; Statistics represent the results of logistic regression analysis.

ADHD symptoms and DRD4 (Table 6). Mother ratings were not significantly associated with DRD4 (Table 6). Teacher ratings were significantly associated with DRD4 when analyzed alone, Wald χ²(N = 90) = 5.28, p = .007, Nagelkerke R² = .088, whereas mother ratings were not, Wald χ²(N = 88) = 0.04, p = .578, Nagelkerke R² = .001. Further, teacher ratings continued to show incremental evidence for association after controlling for mother ratings, Wald χ²(N = 85) = 4.93, p = .013, ΔR² = .086. It was concluded that the best-fitting model, however, included the interaction between mother and teacher ratings. The improvement in fit over and above the main effects model was significant, Wald χ²(N = 85) = 4.34, p = .036, and resulted in a substantial increase in effect size (ΔR² = .082). A visual inspection of this interaction (see Figure 2) suggested that teachers’ ratings showed an increasing linear relation with the number of DRD4 high-risk alleles as mother ratings of these symptoms decreased. This is indicated by the decrease in slopes of the regression lines, which were graphed with mother ratings set at −1 standard deviation, at the mean, and at +1 standard deviation. Because only 2 participants possessed two copies of the DRD4 high-risk allele, this analysis was rerun excluding those participants to ensure they were not solely responsible for the significant interaction. The results remained highly similar after excluding these participants: interaction term, Wald χ²(N = 83) = 4.41, p = .036.

Table 6
Association Analyses Between DRD4 and Mother and Teacher Ratings of Inattentive Symptoms Using Ordinal Regression

<table>
<thead>
<tr>
<th>Model</th>
<th>Overall model</th>
<th>Individual variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>χ²</td>
</tr>
<tr>
<td>M only</td>
<td>88</td>
<td>1.049</td>
</tr>
<tr>
<td>T only</td>
<td>90</td>
<td>1.804</td>
</tr>
<tr>
<td>M &amp; T</td>
<td>85</td>
<td>5.687</td>
</tr>
<tr>
<td>T</td>
<td>88</td>
<td>1.781</td>
</tr>
<tr>
<td>M &amp; T &amp; M’T</td>
<td>85</td>
<td>11.408</td>
</tr>
<tr>
<td>T</td>
<td>85</td>
<td>12.718</td>
</tr>
<tr>
<td>M’T</td>
<td>85</td>
<td>.478</td>
</tr>
</tbody>
</table>

Note. Boldface indicates significance at the .05 level.Italic represents a trend at the .10 level. The overall model p is two tailed. The individual variables p is one tailed, except where noted. For univariate regression analyses, individual variables R² represents the effect size for that model. For multiple regression analyses, ΔR² represents the incremental contribution of the individual variable over and above the other variables in the model; OR = odds ratio; M = mother; T = teacher. DRD4 = dopamine D4 receptor gene.

a Two tailed p. b Assumption of equal slope coefficients across categories was violated; Statistics represent the results of logistic regression analysis.
This best-fitting optimal informant model was then compared with each of the simple algorithms (Table 6). Of these algorithms, only the or rule yielded evidence for a trend toward association between DRD4 and inattentive symptoms, Wald $\chi^2(N = 93) = 2.26, p = .066$, Nagelkerke $R^2 = .038$. Because the test of parallel lines was violated, $\chi^2(1, N = 93) = 4.99, p = .026$, for the ordinal regression analysis evaluating the or rule, this algorithm was evaluated with binary logistic regression. Nonetheless, the optimal informant model that included the interaction between mother and teacher ratings yielded an effect size (Nagelkerke $R^2 = .169$) that was 4 times as large as that of the or algorithm (Nagelkerke $R^2 = .038$). Thus, the optimal informant method appeared to yield stronger evidence for association between inattentive symptoms and DRD4 than the simple algorithms.

Discussion

Candidate gene studies of ADHD frequently use clinicians’ best estimate diagnoses or simple algorithms, such as the or and and rules, to combine data from multiple informants and construct phenotypes for studying linkage and association, but these methods have never been empirically validated or contrasted within the context of a molecular genetic study. In the current study, we systematically evaluated several methods of operationalizing ADHD phenotypes by examining their association with two candidate genes, DAT1 and DRD4. Analyzing symptom scores from multiple informants provided two lines of evidence suggesting that a regression-based optimal informant approach yields more informative phenotypes than do those produced by simple algorithms such as the or, and, and averaging methods.

First, the regression-based optimal informant method allowed for the empirical evaluation of the data provided by each rater. A significant association between DAT1 and teacher ratings of hyperactive–impulsive ADHD symptoms, which yielded an effect size of $R^2 = .045$, was strengthened when mother ratings and the interaction between teacher and mother ratings were added to the regression model, as indicated by a substantial increase in effect size ($R^2 = .119$). In contrast to this interaction model, the association between DAT1 and inattentive ADHD symptoms was sufficiently estimated by teacher ratings alone, yielding an effect size of $R^2 = .100$. For DRD4, a model including the interaction between mother and teacher ratings provided the strongest evidence for association with inattentive symptoms (Nagelkerke $R^2 = .169$), but no relation was detected between DRD4 and the hyperactive–impulsive symptoms.

Taken together, these results serve to highlight the potential complexities that underlie the relations between informant ratings. For example, the interaction between mother and teacher ratings of hyperactive–impulsive symptoms in predicting DAT1 genotype suggested that the strength of the association between teacher ratings and DAT1 increased as mother ratings increased. In contrast, an examination of the interaction between mother and teacher ratings of inattentive symptoms in predicting DRD4 genotype suggested that the strength of the association between teacher ratings and DRD4 increased as mother ratings decreased. The former finding provides an example of how informant ratings might act synergistically to provide evidence for genetic association, whereas the latter result might be indicative of key differences in how parents and teachers rate inattentive symptoms. Support for such an interpretation comes from previous research suggesting that teachers show greater sensitivity than do mothers when assessing inattention in children (Gadow et al., 2004). Thus, these results demonstrate the usefulness of a regression-based approach to evaluate the relations between informant ratings when testing for association between candidate genes and ADHD.

The second line of evidence suggesting that a regression-based optimal informant approach yields more informative phenotypes than do simple algorithms comes from the stronger and more consistent evidence for association produced by the former approach. For example, the regression-based approach yielded stronger evidence for association with DAT1 than did each of the simple
algorithms with effect sizes (i.e., Nagelkerke $R^2$’s) that were two- to threefold larger. Further, this evidence was consistent with previous research in most cases. More specifically, the regression-based approach showed stronger evidence of association between $DRD4$ and inattentive symptoms than did hyperactive–impulsive symptoms, which agrees with previous reports (Lasky-Su et al., 2007; McCracken et al., 2000; Rowe et al., 1998), whereas the simple algorithms failed to yield significant evidence for such an association. Further, the regression-based approach yielded strong evidence suggesting an association between $DAT1$ and hyperactive–impulsive symptoms as well as inattentive symptoms. Though this latter finding contradicts some previous research from our own lab suggesting that $DAT1$ is more highly associated with hyperactive–impulsive symptoms than with inattentive symptoms (Waldman et al., 1998), it is consistent with previous research suggesting a relation between the 10-repeat allele of $DAT1$ and increased risk for ADHD. Thus, it will be of interest to establish whether this finding can be replicated in future studies.

These results have direct implications for how researchers assess symptoms of childhood psychiatric disorders, such as ADHD, for genetic analysis. For example, parent ratings did not show significant evidence for association between ADHD and $DAT1$ or $DRD4$ when analyzed alone, whereas teacher ratings yielded significant evidence for an association between hyperactive–impulsive and inattentive symptoms and $DAT1$ and between inattentive symptoms and $DRD4$. This result clearly demonstrates that teacher ratings, independent of parent ratings, can be highly informative for establishing association between candidate genes and ADHD. Nonetheless, the negative results produced by parent ratings were surprising given the demonstrated validity of both parent and teacher ratings as a measure of a child’s externalizing behaviors (e.g., Bank, Duncan, Patterson, & Reid, 1993; Loeber, Gren, Lahey, & Stouthamer-Loeber, 1991; Verhulst, Koot, & Van der Ende, 1994) and candidate gene studies that yielded evidence for associations between ADHD, $DAT1$, and $DRD4$, with only parent ratings as phenotypes, including studies from our own lab (e.g., Rowe et al., 1998; Waldman et al., 1998). Though surprising, these results indicate that relying on a single informant, such as the mother, may yield a limited perspective on child psychopathology that can reduce a study’s power to detect genetic association. Thus, these findings demonstrate the importance of collecting data from both parents and teachers to assess childhood symptoms of ADHD and suggest that future studies should include a test for informant-specific evidence of genetic association as well as evidence of genetic association derived from the combination of informant reports.

As a further illustration of this conclusion, one potential explanation for the negative findings produced by parent ratings in the current study could be due to the modest sample size. Effect sizes for the relations among $DAT1$, $DRD4$, and ADHD have consistently been small (explaining ~3%–5% of the variance; Faraone and Kahn, 2006; Waldman & Gizer, 2006), and thus, failures to replicate findings for association are common. This could account for the negative findings produced by parent ratings in the present study, and this emphasizes the need for methods that use phenotypic data in a way that maximizes a study’s ability to detect the relation between a candidate gene and a trait or disorder.

The current study also suggests that future molecular genetic studies should carefully consider how to use father ratings. First, the high degree of overlap between mother and father ratings ($r = .72$) in the current study suggests that mother and father ratings may provide largely redundant information for establishing an association between candidate genes and ADHD. Second, the low rate of father participation makes it difficult to conduct tests of association with sufficient statistical power. As a result, future studies with the regression-based optimal informant approach should carefully evaluate whether father ratings provide sufficient independent information when testing for a genetic association that compensates for the reduction in statistical power resulting from their inclusion.

**Limitations and Conclusion**

Finally, there are some potential limitations of the current study that should be addressed. First, the study’s exploratory nature represents an important limitation. To evaluate several methods for constructing phenotypes, multiple statistical tests were conducted and corrections were not used, given that the primary aim was to systematically compare the extent to which each method provided evidence for association among ADHD symptom ratings, $DAT1$, and $DRD4$. Although, in the current study, we attempted to control for chance findings by making predictions regarding the nature of the expected associations on the basis of previous research, the need for replication in independent samples clearly remains. Second, the current study sample was clinic-referred, and thus, the participants selected represent the upper ends of hyperactive–impulsive and inattentive behavior. It would be interesting to test whether the study results could be extended to the general population, but given that a comparison sample was not available, this remains an empirical question for future study. As a third caveat, it should be noted that although the regression-based approach proved useful in the current research context, it is difficult to relate these findings to studies investigating the use of multiple informants in the assessment of ADHD. Thus, the findings of the present study are unlikely to exhibit usefulness in clinical assessment.

These limitations aside, the current study demonstrates the importance of carefully constructing phenotypes for molecular genetic studies. Because the effect sizes for the relations among $DAT1$, $DRD4$, and ADHD in prior research have been relatively small (explaining ~3%–5% of the variance), failures to replicate findings for association are common. Thus, it is critical that such studies use phenotypic data in a way that maximizes a study’s ability to detect the relation between a candidate gene and a trait or disorder. The results of the current study suggest that regression-based methods of combining multi-informant data may provide a more powerful technique for using such data than do simple algorithms such as the or, and, and averaging methods.

**References**


Received July 5, 2007
Revision received May 27, 2008
Accepted May 30, 2008