Behavior Problems in Children With Diabetes: Disentangling Possible Scoring Confounds on the Child Behavior Checklist

Clarissa S. Holmes, PhD, Deedrah Respess, MS, Tammy Greer, PhD, and Johnette Frentz, MD
1Georgetown University, 2Tulane University, and 3University of Southern Mississippi

Objective: The Child Behavior Checklist (CBCL; T. M. Achenbach, 1991), when used to assess the behavior of children with diabetes, may contain confounds because some behavioral items can have a physiologic etiology, and may skew reports of behavioral disturbance.

Methods: Two techniques were used to disentangle possible scoring confounds in the behavioral ratings of children with and without diabetes: (1) the Somatic Complaints scale was deleted, or (2) Diabetes Items, identified a priori with 89% agreement by nine medical personnel, were deleted.

Results: As expected, with traditionally scored protocols, children with diabetes obtained higher Internalizing and Total Behavior Problem scores than controls. This group difference persisted whether the Somatic Complaints scale or the Diabetes Items were deleted.

Conclusions: Compared to controls, children with diabetes obtained mildly elevated scores on six of the eight CBCL scales, regardless of scoring method, suggesting that their mildly elevated behavioral profile is not confounded by physiologic symptomatology.

Key words: diabetes; Child Behavior Checklist; chronic illness.

Numerous studies have demonstrated that chronically ill children are at an increased risk for psychological difficulties (Breslau, 1985; Cadman, Boyle, Scatmari, & Of ford, 1987). For example, Cadman et al. (1987) found that children with chronic illnesses experience a twofold increase in psychiatric disorders compared to healthy, unaffected peers. Lavigne and Faier-Routman (1992) conducted a meta-analysis of studies of children's psychological adjustment to 20 different pediatric disorders. Children with chronic illnesses were found to have both internalizing or overcontrolled behaviors, as well as externalizing or undercontrolled symptomatology, that was between a third to a half of a standard deviation above the norm. Although the average metaanalytic effect size for children with diabetes was .58, studies of children with diabetes exhibited the widest range of effects from -.2 to 1.4, compared to other pediatric disorders.

Some of the outcome variability in studies of diabetes may stem from the use of different assessment measures. The advent of well-developed and normed behavioral checklists, such as the Child Behavior Checklist (CBCL; Achenbach & Edelbrock, 1983), has provided a standardized method of obtaining and categorizing information about chil-
dren's behavior. For example, Wertlieb, Hauser, and Jacobson (1986) obtained behavioral ratings on the CBCL that indicated children with diabetes experienced significantly higher internalizing scores than an acutely ill control group. Northam, Bowden, Anderson, and Court (1992) analyzed their CBCL data by the percentage of children scoring in the clinical range. Significantly more children with diabetes had Internalizing Behavior Problems (19%) and Total Behavior Problems (27.8%) than children from the test normative data.

Perrin, Stein, and Drotar (1991) have argued that the CBCL Somatic Complaints scale, which is part of the Internalizing Behavior Problem Score, contains considerable bias for children with chronic illnesses. The potential confound from these nine physical symptom items may positively skew reports of internalizing symptoms. They note that the CBCL was not created for or recommended for use with medically ill children, despite its widespread use. For example, instructions for the Somatic Complaints items are to check only "physical symptoms without medical cause" such as "headaches, nausea, or problems with eyes," but it is not known to what extent parents can factor out the portion of a behavior that may have a physical cause but also may have a psychogenic component as well.

In a series of CBCL studies, Wallander and associates (Wallander, Varni, Babani, Banis, & Wilcox, 1988; Wallander, Varni, Babani, Banis, DeHaan, & Wilcox, 1989) examined this problem by deleting the Somatic Complaints scale and calculating a modified Internalizing Behavior Problem T-score. When children with six different pediatric disorders were evaluated (Wallander et al., 1988), deletion of the Somatic Complaints scale yielded Internalizing Behavior Scores that dropped 2 points but still remained higher than the test norms. More dramatic results were found in another study (Wallander et al., 1989) when only physically handicapped children were evaluated. After deletion of the Somatic Complaints scale, Internalizing Behavior Problem scores were lowered a substantial 7.5 points from the clinically significant range to within the clinically "normal" range, although scores again remained significantly higher than the norm.

In addition to the Somatic Complaints scale, physical symptoms items are found on other CBCL scales (Perrin et al., 1991). For example, the item "clumsy," which is found on the Social Problems scale, may be a symptom of hypoglycemia. Thus, there remains a possibility of obtaining inflated estimates of behavioral disturbance with pediatric populations from other disease-sensitive items embedded throughout the checklist. To remove all potentially confounded physical symptoms from the CBCL, it would be necessary to first determine all disease-sensitive items and then to remove them, regardless of their location.

The proposed research was designed to investigate two scoring adaptations of the CBCL for children with diabetes and a matched control group. First, the Somatic Complaints scale was deleted from the determination of Internalizing Behavior Problems and Total Behavior Problems to evaluate if scores were differentially lowered for a chronic illness group compared to a control group. Next, diabetes-related CBCL items, identified by pediatric endocrinology personnel, were used to devise a Diabetes Item scale to determine if it was more sensitive in removing possible disease-related confounds from estimates of children's behavioral functioning.

Methods

Subjects
One hundred thirteen children with IDDM, ages 8 to 16, (M = 11.7 years, SD = 2.4) participated in this study. Children were recruited from several diabetes clinics in a metropolitan area as part of a larger ongoing study. At the time of testing, all children with diabetes were at least six months postdiagnosis and free of secondary disease complications (i.e., retinopathy, nephropathy, etc.) and other major medical diseases. Children had an average age of disease onset of 8.3 years (SD = 3.1) and an average disease duration of 3.2 years (SD = 2.6). With a nondiabetic reference range of <5.5%, children had an average Hba1c of 11.7% (SD = 3.0). On average, children were from the middle socioeconomic status (SES) strata (M SES raw score = 41.2, SD = 15.9), according to the Hollingshead Four Factor Scale (Hollingshead, 1975).

The children comprising the control group (N = 83) were recruited from parochial and public schools within the same metropolitan area. They were matched in group composition with the children who had diabetes for grade, race, and gender. Controls were similar to children with diabetes in age (M = 11.7, SD = 2.5) and SES (M SES = 43.7, SD = 13.9). Children serving as controls also were free from major illness or trauma.
Procedure
Children with diabetes and their parents were contacted by mail for study enrollment, with a follow-up telephone call to answer questions. Testing occurred on the same day or within two weeks of children's scheduled medical appointment. Parents were administered the CBCL while their children were being tested. Control children were recruited from the enrollment roster of parochial and public schools within a metropolitan area. For each grade level, children's names were masked and a match for gender and race was randomly selected for a target child with diabetes.

Seven endocrinologists and three nurses who specialized in pediatric diabetes care from several metropolitan hospitals were contacted by letter to solicit their assistance in rating disease-related items from the CBCL. Nine of the 10 responded with item ratings. Five of the pediatric endocrinologists who responded were affiliated with one of two different university hospitals, and the sixth was based at a large private hospital. The three diabetes nurses were located one at each of the three hospitals. Medical raters were given the 118 behavior problem items from the CBCL and asked to note for each item those which might be diabetes sensitive with the following scale: 0 = not true of or no association with diabetes; 1 = somewhat true of or sometimes associated with diabetes; and 2 = very true of or often associated with diabetes.

Measures
The Child Behavior Checklist (CBCL; Achenbach, 1991). The parent-completed CBCL consists of 118 behavior items designed to measure the behavioral problems of children ages 4 through 18. The 1991 revision of the CBCL consists of the same 118 behavior items as the 1983 version (Achenbach & Edelbrock, 1983), with scoring changes made and the provision of new national norms. In contrast to the age- and gender-specific scales of the 1983 checklist, the 1991 edition provides a single set of behavior scales across groups, making direct behavioral comparisons possible across different groups of children. Because the Sex Problem scale only occurs in the 6-11 age range, and is not scored as part of the Internalizing score, it is not reported here.

Scoring Methods
Several scoring methods were proposed for analysis in the present study: Traditional, Somatic Complaints deleted, and Diabetes Items deleted. These scoring procedures were conducted as follows:

Traditional Scoring. Traditionally scored (i.e., not modified) answer sheets of children with and without diabetes were compared.

Deletion of Somatic Complaints scale. All items comprising the Somatic Complaints scale were rescored as zero, thereby deleting the scale.

Deletion of Diabetes Items. CBCL items selected by pediatric endocrinologists and nurses as being diabetes-sensitive were rescored as zero, effectively deleting the items. A total of seven Diabetes Items were identified as disease-sensitive by eight of the nine medical raters, for an endorsement rate of 89%. At the 89% level of agreement, 6% of the CBCL behavior problem items were deleted. When a less stringent selection criterion of 78% agreement was used, a total of 18 items (15%) were endorsed as diabetes-sensitive. Because the construct validity of the CBCL may be jeopardized with the removal of 15% of the items, and because the deletion of seven items was comparable in number to the deletion of nine items from the Somatic Complaints scale to allow a fair comparison of relative item sensitivity, the criterion of 89% agreement was used for the purposes of the analyses that follow. Table I lists

<table>
<thead>
<tr>
<th>Diabetes items</th>
<th>78% Endorsed</th>
<th>89% Endorsed</th>
<th>Somatic complaint items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feels dizzy</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Overeating</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Overweight</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Headaches</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Stares blankly</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Sudden changes in mood or feeling</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Wets the bed</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Overtired</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Nausea, feels sick</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Problems with eyes</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Stomach aches or cramps</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Vomiting, throwing up</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Can't concentrate,</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>or pay attention for long</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confused or seems to be in a fog</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Doesn't eat well</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Nightmares</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Stubborn, sullen, or irritable</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Wets self during the day</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>
the Diabetes Items endorsed by 89% and 78% of medical personnel. It also shows which Diabetes Items are found on the Somatic Complaints scale according to each criterion. Because only two of the seven Diabetes Items identified by medical raters at the 89% level of agreement also were included on the Somatic Complaints scale, these scales had only marginal overlap.

Results

Analyses of variance (ANOVAs) were conducted to determine if children with and without diabetes differed according to scoring technique on the CBCL. These same three scoring methods were examined with a subsample of subjects who obtained Total Behavior Problem scores in the clinically elevated range ($T \geq 70$). The alpha error level for all statistical tests, including the Newman-Keuls post hoc analyses, was set at .05.

Traditional Scoring Method

A multivariate analysis of variance (MANOVA) was conducted with the following dependent variables: Internalizing, Externalizing, and Total Behavior Problems scores. Disease status (diabetes/control) was the independent variable. The main effect of disease status was significant, $F(3,192) = 5.67$, $p < .05$. Children with diabetes exhibited significantly more Internalizing, $F(1,194) = 10.97$, $p < .05$, Externalizing, $F(1,194) = 15.70$, $p < .05$, and Total Behavior Problems than control children, $F(1,194) = 16.78$, $p < .05$. The $M$s and $SD$s for the Internalizing, Externalizing, and Total Behavior Problem scores of subjects are presented in Table II.

Percentage of Clinically Elevated Profiles

Because it was expected that more children with diabetes would have clinically elevated Total Behavior Problem scores than control children, the difference in the proportions of subjects with and without diabetes who had clinically elevated Total Behavior Problem scores ($T \geq 70$) was computed. Of children with scores in the clinically elevated range of $T \geq 70$ ($N = 12$), 83.3% were children with diabetes and 16.7% were controls ($n = 10$ vs. $n = 2$, respectively). Using the Exact Binomial Probability, the probability of an outcome this or more extreme was significant ($p = .039$). Evaluated another way, approximately 9% (10 of 113) of the sample with diabetes exhibited clinically elevated scores versus only 2% (2 of 83) of control children.

Deletion of Somatic Complaints

A modified Internalizing Behavior Problem score was computed for subjects with and without diabetes by deleting the Somatic Complaints scale. A $2 \times 2$ repeated measures ANOVA was conducted with scoring method (traditional/somatic complaints deleted) as the repeated measure, and disease status (diabetes/control) as the independent variable for the Internalizing Behavior Problem scores. No interaction was found, but there was a main effect of scoring method, $F(1,194) = 114.99$, $p < .05$. As expected with the deletion of nine items, the Internalizing scores resulting from the removal of the Somatic Complaints scale were significantly lower than traditional Internalizing scores ($M = 50.3$, vs. $M = 52.0$, respectively) for the combined groups of
children. A main effect of disease status also was found to be significant, \( F(1,194) = 11.03, p < .05 \). Subjects with diabetes continued to exhibit significantly higher Internalizing scores relative to controls on the modified Internalizing score. That is, deletion of the Somatic Complaints scale lowered the Internalizing scores for both children with and without diabetes (\( M = 50.8, \) vs. \( M = 49.0, \) respectively), although children with diabetes continued to score significantly higher than controls. Failure to find a significant interaction suggests that children with diabetes did not experience a differential lowering of their Internalizing scores with the removal of these physiologic behavioral items.

Next, the impact of removing the Somatic Complaints scale was assessed for Total Behavior Problem scores with a repeated measures ANOVA. Similar to Internalizing scores, a main effect of scoring method was found, \( F(1,194) = 139.79, p < .05 \), along with a main effect of disease status, \( F(1,194) = 16.71, p < .05 \). There was no interaction of the variables. Deletion of the Somatic Complaints scale yielded Total Behavior Problem scores for the combined study groups that decreased when compared to the traditionally calculated scores (\( M = 50.1, \) vs. \( M = 51.7, \) respectively). Children with diabetes continued to obtain significantly higher Total Behavior Problem scores than their controls. See Table III for Ms and SDs.

Deletion of Diabetes Items
Next, Diabetes Items were deleted and Total Behavior Problems were evaluated with a repeated measures ANOVA. A main effect of type of scoring method (traditional/diabetes items deleted), was found, \( F(1,193) = 184.81, p < .05 \). Not surprisingly, the Total Behavior Problem scores of the combined study groups were significantly lower with the deletion of the seven diabetes items compared to the traditionally calculated scores (\( M = 49.9, \) vs. \( M = 51.7, \) respectively). Children with diabetes continued to obtain significantly higher Total Behavior Problem scores than their controls. See Table III for Ms and SDs.

Comparison of the Three Scoring Methods
In an effort to investigate differences in Total Behavior Problem scores among the various scoring procedures, a \( 2 \times 3 \) repeated measures ANOVA was conducted with disease status as the independent variable and scoring method (Traditional/Somatic Complaints deleted/Diabetes Items deleted) as the repeated measure. Main effects of disease status, \( F(1,193) = 16.74, p < .05, \) and scoring method, \( F(3,191) = 78.95, p < .05, \) were significant.

A Newman-Keuls range test revealed that the Total Behavior Problem scores of the combined study groups (diabetes/control) were significantly higher with the traditional scoring method (\( M = 51.7 \)) compared to the alternative scoring methods: Deletion of the Somatic Complaints scale (\( M = 49.9 \)), and Deletion of the Diabetes Items (\( M = 50.1 \)). Deletion of the Somatic Complaints scale, however, produced significantly lower Total Behavior Problem scores than Deletion of the Diabetes Items. Failure to find a significant interaction indicates that none of the proposed scoring techniques differentially affected the scores of children with diabetes compared to those of controls.

Clinically Elevated Profiles and Scoring Methods
To further investigate differences in the scoring methods, analyses were conducted for those subjects (\( N = 12 \)) who obtained Total Behavior Problem scores in the clinical range (\( T \geq 70 \)). No differences were found in the pattern of Internalizing, Externalizing, and Total Behavior Problem scores from those that were obtained for the total sample. Because of redundancy, these analyses are omitted.

Discussion
The goal of this investigation was to determine if physical symptomatology items on the Child Behavior Checklist (CBCL; Achenbach, 1991) might
inadvertently confound estimates of behavioral problems in a group of children with diabetes. Regardless of how the data were analyzed to address this question, the consistent answer was that the behavioral scores of children with diabetes were not differentially affected by potentially confounded physical behavioral symptomatology (e.g., "feels dizzy") compared to matched controls.

Although children with diabetes obtained significantly higher CBCL Internalizing, Externalizing, and Total Behavior Problem scores compared to controls, consistent with the literature (Northam et al., 1992; Wertlieb et al., 1986), this apparently was not due to a possible confound from the nine physical symptoms on the Somatic Complaints scale. When the CBCL Somatic Complaints scale was deleted, Internalizing and Total Behavior Problem scores were lowered similarly for both children with and without diabetes, by about 1 point, consistent with previous findings (Wallander et al.; 1988). Failure to find a differential lowering of group scores indicates that physiologic items are probably not skewing diabetes-related findings of higher Somatic Complaints and Internalizing scores that have been reported here and elsewhere (Northam et al., 1992, Wallander et al., 1989; Wertlieb et al., 1986). Thus, deletion of the Somatic Complaints scale to obtain "confound-free" scores is unwarranted and not clinically useful.

In the present study, children with diabetes evidenced their highest scale score on the Somatic Complaints scale, congruent with the findings of others (Wallander et al. 1988); however, surprisingly, the control children also obtained a high point on this scale, underscoring the importance of control comparisons. It may be that the ambiguous, almost "flu-like" symptoms of the Somatic Complaints scale (i.e., headaches, dizziness, stomach aches) are readily endorsed by parents of a majority of children, especially those from the middle classes, as in the present study. Even though both groups of children had high points on the Somatic Complaints scale, children with diabetes still obtained significantly higher Somatic Complaints scores than controls, although both means were still in the average range (M = 58.7 and 56.0, respectively). Equally surprisingly, this scale difference did not discriminate the behavioral profiles of groups of children. Children with diabetes scored significantly higher than controls on all of the CBCL scales except two, the Social Problems and Attention Problems scales. This pattern of general global scale elevation suggests that diabetes is primarily associated with general mild, nonspecific behavioral symptomatology in the average range, and not a specific pattern of behavioral functioning.

Because the diffuse physical symptoms on the Somatic Complaints scale may be relatively insensitive to the specific symptoms associated with diabetes, a second scoring method was devised based upon an a priori symptom identification of seven diabetes-sensitive behavioral items by pediatric endocrinology specialists. See Table I for the Diabetes Items. One could convincingly argue that these items should not be included as indicators of behavior problems for children with diabetes. However, deletion of the Diabetes Items produced similar decreases in Total Behavior Problems scores, of approximately 1 point, for children with and without diabetes. Interestingly, when the two scoring methods were compared, deletion of the Somatic Complaints scale produced an overall significantly greater decline in scores than the deletion of the Diabetes Items. However, given the failure to find an interaction between disease status and scoring method, the most parsimonious explanation for this finding is that removal of nine items (Somatic Complaints) lowers scores more than removal of seven items (Diabetes Items). Thus, although both scoring adaptations possessed content validity, comparisons between study groups indicated that neither technique had discriminant validity and, therefore, no clinical utility.

Although the Diabetes Items were selected by medical experts on the basis of their clinical experience, only 5% of the entire sample obtained scores on these items. This datum, along with only mildly elevated group behavioral scores that were well within the average range, in this and other studies (Northam et al., 1992, Wallander et al., 1989; Wertlieb et al., 1986), suggests there is little benefit to be gained from further exploration of these CBCL scoring adaptations with children who have diabetes. As Perrin et al. (1991) note, the CBCL was developed to identify psychopathology, not to describe children with subtle or mild behavioral symptoms. In addition, even when a more behaviorally disturbed subsample of children with and without diabetes was evaluated (i.e., who obtained scores in the clinically elevated range), the scoring adaptations produced no differential group results, further reiterating the limited utility of exploring CBCL scoring adaptations in this population.

A potential limitation to this study is that results were based on the 1991 version of the CBCL (Achenbach, 1991), and not the 1983 version...
Behavior Problems in Children With Diabetes

(Achenbach & Edelbrock, 1983), which was used to assess the behavioral status of children in the earlier literature cited. There were several scoring changes accompanying the 1991 version. Nevertheless, despite these scoring differences, findings and interpretations from the present study are remarkably similar to those reported in the past with the 1983 version (i.e., a 1- to 2-point drop in Internalizing scores; Wallander et al., 1988). Further, it is important to remember that although some items changed scales between the 1983 and 1991 versions, the same 118 behavioral items comprise each version.

Finally, previous research has shown that physically impaired populations, such as children with spina bifida and cerebral palsy, may experience significantly higher rates of behavioral problems than children with diabetes (Wallander et al., 1989), as well as a greater drop (7.5 points) in Internalizing scores when the Somatic Complaints scale from the CBCL is deleted. Future efforts to devise disease-specific scoring modifications to the CBCL may be better directed toward children with illnesses such as those described that involve physical sequelae and/or a higher incidence of significant behavior problems.

Acknowledgments

This work was based on a thesis by the second author in partial fulfillment of a Master's of Science degree and was funded in part by grant DK37545 to the first author. We thank Drs. Grace Banuchi, Jayasharee Rao, Alfonso Vargas, Martin Young, and Teresa Zimmerman for their valuable assistance rating disease items; JoAnne DeLaune, Terry Compton, and Trudy Parker for their assistance with rating and with scheduling patients for evaluation; the Orleans Parish public schools and the St. Stevens, Mercy Academy, De La Salle, Holy Name, and Mount Carmel schools in the New Orleans area for their cooperation with this project.

Received April 8, 1997; accepted September 16, 1997

References


*Nine items scored as Internalizing problems on the 1983 version of the CBCL were switched to the Externalizing scale in the 1991 version, although only one of these was a Diabetes Item ("sudden changes in mood or feeling"); and this item changed only for boys ages 12-16 years. Eight items from the 1983 Externalizing scale were moved to the Internalizing scale on the 1991 version of the checklist, although only one of these items was a Diabetes item ("stares blankly"), and it switched only for girls ages 6-11 years.