A Family-Based Eating Disorder Day Treatment Program for Youth: Examining the Clinical and Statistical Significance of Short-Term Treatment Outcomes

KATHERINE HENDERSON
Department of Psychology, Carleton University; and Anchor Psychological Services Inc., Ottawa, Ontario, Canada

ANNICK BUCHHOLZ
Department of Psychology, Carleton University; and Centre for Healthy Active Living, Children’s Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada

NICOLE OBEID
Child and Youth Mental Health Research Laboratory, University of Ottawa; and Eating Disorders Program, Children’s Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada

ANNIK MOSSIERE
Department of Psychology, University of Saskatchewan, Saskatoon, Saskatchewan; and Eating Disorders Program, Children’s Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada

DANIJELA MARAS
Department of Psychology, Carleton University; and Eating Disorders Program, Children’s Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada

MARK NORRIS, MEGAN HARRISON, and STEPHEN FEDER
Department of Pediatrics, University of Ottawa; and Eating Disorders Program, Children’s Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada

WENDY SPETTIGUE
Department of Psychiatry, University of Ottawa; and Eating Disorders Program, Children’s Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada

This article describes an innovative family-based day treatment program (DTP) for youth with moderate to severe eating disorders.
A sample of 65 youth completed a battery of psychological measures pre- and post-treatment and 6 months after program completion. Treatment outcomes were assessed in three main domains: (a) medical stabilization, (b) normalization of eating behavior, and (c) improved psychological functioning. Overall, patients demonstrated statistically significant and clinically meaningful improvements on all outcome measures. Findings indicate that a comprehensive DTP can successfully facilitate positive outcomes in youth with eating disorders and that these improvements can be maintained 6 months post-treatment.

Anorexia nervosa (AN), bulimia nervosa (BN), and eating disorder not otherwise specified (EDNOS) are serious health conditions that affect children and adolescents (Ackard, Fulkerson, & Neumark-Sztainer, 2011; Silber, 2005; Wilfley, Kass, Kolko, & Stein, 2011), with prevalence rates estimated as high as 1.5% in young females (Hoek, 2006; Swanson, Crow, Le Grange, Swendsen, & Merikangas, 2011). These eating disorders (ED) are associated with premature mortality and high morbidity (Gowers & Bryant-Waugh, 2004; Smink, van Hoeken, & Hoek, 2012). Although youth are recognized as a serious at-risk group for EDs (many adult EDs develop during adolescence), relatively little research has been dedicated to this group in terms of treatment outcomes (Gowers & Bryant-Waugh, 2004; Lock, 2010).

Treatment interventions for patients with EDs traditionally have been offered on an outpatient or inpatient basis, but this has shifted since the recent introduction of day treatment programs (DTPs) in the last few decades (Lammers, Exterkate, & De Jong, 2007; Ogrodniczuk & Steinberg, 2005; Willinge, Touyz, & Thornton, 2010; Zipfel et al., 2002). Day treatment programs were originally designed to involve intensive psychological, social, and biological interventions (Piran, Langdon, Kaplan, & Garfinkel, 1989). The therapeutic goals are to normalize eating and related behaviors, achieve medical stabilization, and improve psychological functioning. Day programs follow a set schedule and operate in a structured format, deemed essential in providing a solid therapeutic environment (Piran et al., 1989). While DTPs are becoming more common in the management of EDs over the last 20 years, there are few studies examining the outcomes of such programs, and their effectiveness in meeting treatment goals (Lammers et al., 2007).

In addition, the current literature on outcomes and program evaluation of DTPs is limited to adult studies, and generally has been characterized by small sample sizes (Fittig, Jacobi, Backmund, Gerlinghoff, & Wittchen, 2008), thus limiting the generalizability of the findings. Of the studies that have been performed on DTPs in adults, the sample sizes have ranged from as low as 12 (Birchall, Palmer, Waine, Gadsby, & Gatward, 2002) up to 110 participants (Peake, Limbert, & Whitehead, 2005), making it clear that further
studies with large sample sizes are needed. Despite this methodological limitation, overall, the few studies that have evaluated DTPs for EDs generally have reported good short-term outcomes (Ben-Porath, Wisniewski, & Warren, 2010; Dancyger, Fornari, & Schneider, 2003; Fittig et al., 2008; Gerlinghoff, Backmund, & Franzen, 1998; Kaplan & Olmstead, 1997; Kong, 2005; Lammers et al., 2007; Piran et al., 1989; Thornton, Beumont, & Touyz, 2002; Willinge et al., 2010; Zeeck, Herzog, & Hartmann, 2004), demonstrating that DTPs are an effective treatment option for adults with moderate to severe EDs.

A further consideration when evaluating clinical program outcomes is to examine different forms of change. Jacobson, Roberts, Berns, and McGlinchey (1999) have suggested the concept of clinical significance for measuring change in the field of mental health. This concept involves two criteria: (a) the magnitude of the change has to be statistically reliable, and (b) by the end of therapy or program, recovered patients should end up in a range that renders them indistinguishable from the well-functioning people (Jacobson et al., 1999). Clinical significance is determined using Reliable Change Index (RCI) analyses of outcome data. This approach allows researchers and clinicians to examine changes in medical and psychological health in a more concrete and clinically relevant manner (compared to using effect sizes, for example). RCI methods examine scores at the individual level, whereas statistical analysis of group data, such as traditional $p$ values and effect sizes, only inform researchers about the overall effectiveness of a treatment. RCI analyses can assess variability in treatment outcome that statistical comparisons between group means cannot (Ben-Porath et al., 2010). Ben-Porath et al. (2010) is the only study to date that has examined outcomes of an ED DTP using clinical significance to evaluate change. Findings from this study indicated statistically significant reduction on all eating disordered outcomes in their adult population. With respect to the clinical significance testing, analyses revealed that the majority of the patients made clinically significant and reliable change by the completion of treatment on all ED measures. By using RCI’s clinical significance, researchers can make more meaningful comparisons and interpretations of data sets and treatment approaches in ways that are not possible with traditional statistical significance testing (Ben-Porath et al., 2010). A standard statistical significance comparison between groups rarely determines the practical importance of the treatment effects (Jacobson et al., 1999). In applied and clinical fields, this practical importance is paramount. No study to date, in the area of EDs, has investigated the clinical significance of treatment outcomes in a youth population. Thus, the goal of the current study was not only to assess the outcomes of a youth DTP, but also to evaluate the clinical significance and meaningfulness of the outcomes.

There remains a large gap in the literature surrounding our understanding of the effectiveness of treatment strategies for EDs in youth (Robergeau,
Joseph, & Silber, 2006). Literature to date, including the research cited above, has neglected to examine and evaluate the outcomes of youth DTPs for EDs. This gap highlights the uniqueness of the current study. In addition, this article involves a larger sample size than most of the adult studies, which will allow us to generalize to youth populations with various EDs. Further, as it seems likely to expect positive changes in patient health throughout intensive treatment; what is even more imperative to examine is whether changes and/or improvements are maintained upon leaving the structured environment of the DTP. To address this issue, the present study goes beyond evaluating patients’ change from pre to post-program, and includes an additional follow-up period at 6 months post-program.

This article aims to comprehensively evaluate the outcomes of a “Maudsley” family-based (Lock, Le Grange, Agras, & Dare, 2001) youth DTP based on all stipulated short-term goals of treatment, using both clinical and statistical significance in a transdiagnostic sample (Fairburn, Cooper, & Shafran, 2003). It is predicted that participants who complete the DTP will make improvements across the three main treatment goals: (a) medical stabilization (see Birchall et al., 2002; Halvorsen, Andersen, & Heyerdahl, 2004), (b) normalization of disturbed eating behavior (see Zeeck et al., 2004), and (c) improved psychological functioning (see Peake et al., 2005), and that these improvements will be maintained 6 months after program completion (Gerlinghoff et al., 1998).

METHOD

Sample and Program Characteristics

The Eating Disorder Day Treatment Program (EDDTP) is located at a pediatric tertiary care hospital in Canada. The program has a maximum capacity of 8 youth, and operates from 8 a.m. to 6 p.m., 5 days per week, with a typical length of 12–14 weeks. This innovative program establishes an environment in which psychological, medical, and nutritional support is available for youth with moderate to severe EDs.

The program is modeled around a family-based treatment philosophy of care, wherein parents’ guilt about causing the eating disorder is lifted, anxiety about the severity of the illness is raised, and parents are empowered to be involved in their child’s recovery. Each patient receives “Maudsley” (Lock & Fitzpatrick, 2007; le Grange, Lock, Loeb, & Nicholls, 2010) informed family therapy. The Maudsley principles are based on striving to bring about weight restoration and to restore the youth’s developmental trajectory, while using the family as a very active component of the treatment team. Parents and youth have an initial meeting with a dietitian and then meet weekly with a psychologist or psychiatrist for Maudsley informed family therapy. In order to provide consistency between the day treatment program meals
and home meals, patients and families receive a meal plan that is co-created with the dietitian. In family therapy, parents are supported in providing constant meal support/supervision with their youth and actively engaging in symptom interruption (i.e., preventing purging, exercising etc.) through the use of close monitoring and distraction within a framework of love, compassion, and empathy. Following the Maudsley method, parents are reminded that they are the experts on their youth, and as therapy progresses when or if appropriate parents are supporting in handing control for eating back to their youth.

During the course of the program, all patients receive group therapy, meal support, academic support, therapeutic outings, and individual and family therapy. This involves a multidisciplinary treatment team consisting of psychiatrists, psychologists, an adolescent health physician, a nurse, a dietitian, a psychometrist, child and youth counselors, a yoga instructor, and several teachers. This program is often used as a step-down treatment plan for those who successfully complete the inpatient program. Following DTP, youth receive supportive outpatient services for six-months to a year post program completion.

The participants in this study consisted of 65 female youth who had presented for an initial intake assessment of an ED and attended the DTP. Patients were referred to the Eating Disorder Service via their family physician or the hospital’s emergency department. Because very few male youth have been referred to the DTP, and because their clinical profiles tend to differ from that of females (e.g., Anderson & Bulik, 2004; Lewinsohn, Seeley, Moerk, & Striegel-Moore, 2002), males were not included in the present study. All female youth in this sample met criteria for a moderate to severe ED according to the *Diagnostic and Statistical Manual of Mental Disorders, 4th edition* (DSM-IV; American Psychiatric Association, 1994), and as established by a consensus diagnosis. Admission to the DTP was determined for patients between 11 and 18 years of age who required a high level of treatment based on symptomatology.

Procedures

Upon arrival to the Eating Disorder Program for an initial intake assessment of an ED, all individuals were asked to complete a battery of self-report standardized measures evaluating medical, nutritional, and psychological functioning. In addition to these measures, a clinical interview performed by a physician and psychologist/psychiatrist also was conducted in order to reach a consensus diagnosis. The patients’ initial assessment was considered their original pre-treatment results as long as they began the program within 4 weeks of the initial assessment. If treatment began after the 4-week mark, patients were asked to again complete the battery of questionnaires, and these results were then used as their pre-treatment information.
Post-treatment data were gathered at discharge from treatment, and again at 6 months post-discharge. Due to non-responses, incomplete data, and missing follow-ups, some of the results presented below do not include all 65 participants, and as such are based on different sample sizes. The follow-up visit that was closest to 6-months (between 4–7 months) from program discharge was used to assess outcomes. This follow-up period was selected because it was thought that a 6-month duration after program completion would reflect the program’s contribution to a patient’s stabilization and/or recovery since it followed them at least twice as long as the time they spent in the program. This study received approval from the Research Ethics Review Board.

Measures

**Body mass index (BMI).** Body mass index (BMI; kg/m²) was used as a marker of medical status for the purpose of this study. Patients’ weight and height were taken by a clinician at baseline, at post day hospital treatment, and again at 6-month follow-up. BMI is one of the most commonly used measures of medical stabilization. Other popular measures of medical stabilization include blood pressure and heart rate. Given that vital signs typically stabilize earlier in the recovery process (e.g., Shamim, Golden, Arden, Filiberto, & Shenker, 2003), and patients admitted to the DTP have generally achieved vital sign stabilization, BMI was used as the primary indicator of medical stabilization and weight recovery (Loeb, Brown, & Goldstein, 2011).

**Children’s Depression Inventory (CDI).** The Children’s Depression Inventory (CDI; Kovacs, 1992) was used to evaluate symptoms of depression. This measure is designed to assess the severity of depressive symptoms in children aged 7 to 17 years, and is comprised of 27 items. Within each item, the youth chooses the statement that best describes his or her feelings within the last 2 weeks. This measure is sensitive to changes in depression over time and is acceptable as an index of the severity of depressive disorder (Kavan, 1992). The CDI demonstrates excellent internal consistency (Cronbach’s alpha .71–.87) across various populations and treatment settings, and test-retest reliabilities are adequate at .75 to .87 (Kavan, 1992; Kovacs, 1992). Patients’ Total CDI T-scores were used in the analyses.

**Eating Disorder Inventory-2 (EDI-2).** The Eating Disorder Inventory-2 (EDI-2; Garner, Olmstead, & Polivy, 1983) was used to provide a comprehensive assessment of the behavioral and psychological characteristics of EDs. The EDI-2 is a reliable and valid 91-item multidimensional self-report instrument used for patients over the age of 12 years that assesses attitudes, feelings, behavior, food, and eating. On a 6-point Likert scale, individuals indicate how often they engage in the queried characteristics. Scores are then weighted from 0 to 3, with higher scores indicating more severe symptomatology. The EDI-2 is divided into 11 subscales, and has been found to possess good test-retest reliability (Crowther, Lilly, Crawford, & Shepherd, 1990; Wear & Pratz, 1987) and good internal consistency with...
alphas ranging from .80 to .91 (Crowther et al., 1990). Two of the subscales were used for the purpose of this article—Drive for Thinness (DT), and Body Dissatisfaction (BD).

**Multidimensional Anxiety Scale for Children (MASC).** The Multidimensional Anxiety Scale for Children (MASC) is a 39-item self-report questionnaire that evaluates a wide spectrum of common anxiety symptoms in youth aged 8 to 19 years (March, Parker, Sullivan, Stallings, & Conners, 1997). Responses to the items range from 0 to 3, with higher scores indicating more clinical levels of anxiety. This measure also contains a Total Anxiety Scale, which offers an overall score for anxiety, as well as an Anxiety Disorder Index. Internal reliability for this measure is adequate with alphas ranging from .50 to .89, and test-retest reliability has been measured at .69 to .93 (March et al., 1997). The Total Anxiety Scale was used for the present analyses.

**Analyses**

All analyses were conducted using IBM SPSS Statistics Software Version 19. A series of one-way repeated measure analyses of variance (ANOVA) and pairwise comparisons were performed to examine statistical changes on psychological and medical outcome variables across three time-points; pre-treatment, post-treatment, and 6 months post treatment. These repeated measures allowed for statistical observations of maintained effects from programming to 6 month after program completion.

To evaluate whether there were clinically significant changes in the outcome variables from baseline to follow-up (pre to post, and post to 6-month follow-up), the Reliable Change Index (RCI) was calculated. Separate RCI analyses were conducted for the two time-set comparisons: one analysis looking at pre to post, and the other looking at post to 6 month. This measurement allowed for the calculation of difference scores which were then compared to a selected cutoff score that was determined for each variable. Based on the validated clinical cutoffs for the Total CDI and MASC scores, T-scores of 65 were used as the RCI cutoff for these measures. A BMI of 19 kg/m² was used as the cutoff for medical stabilization, where individuals with a BMI under 19 were identified as underweight (Howard, Evans, Quintero-Howard, Bowers, & Arnold, 1999). The cutoffs for the EDI-2 were at the 68th percentile (i.e., one SD from the mean), resulting in a cutoff score of 10 for the DT subscale, and a cutoff score of 18 for the BD subscale (see Kordy, Percevic, & Martinovich, 2001).

Jacobson and colleagues (1999) state that the RCI statistical procedure is a reliable method to evaluate the clinical changes in patients’ symptomatology over time. “Reliable” change represents change of more than 2 standard deviations.

---

1 RCI Computational Formula for Individual Data (Jacobson & Truax, 1991); $RCI = \frac{x_{pre} - x_{post}}{S_{diff}}$, $S_{diff} = (\frac{2(SE)^2}{1})^{1/2}$; $SE = SD_{pre} (1 - r_{xx})^{1/2}$
errors (SE) of measurement for individuals who needed to change (i.e., those whose scores were above the normal cutoff for t-scores or below the cutoff for BMI). Several types of change are calculated with RCI. For this study, change was coded into four categories: unchanged, deteriorated, improved, and irrelevant change. Irrelevant change refers to individuals who started off in the normal range on a particular scale, making their positive change irrelevant to the treatment goals of these particular individuals (Jacobson & Truax, 1991). For example, patients within a healthy BMI range at pre-treatment who remain within a healthy BMI range post-treatment and at follow-up would be classified in irrelevant change category. Within the improved category, a subgroup was classified as recovered (clinically significant change) for those who qualified as reliably improved with a post treatment or follow-up score that had moved into the normal range. RCI analyses do not provide an additive effect to see how changes are maintained pre-post; rather they identify how many people moved into the recovered range during the program and how many moved into the recovered range after leaving the structured program.

RESULTS

Participants ranged in age from 11 to 17 years, with a mean of 15.0 years (SD = 1.34). In terms of diagnosis, 63.7% presented with AN, 10.2% with BN, and 26.1% with EDNOS. The participants’ BMI (kg/m²) at baseline ranged from 13.37 to 26.00, with a mean of 18.71 kg/m² (SD = 2.40). The average length of treatment was 14.79 weeks (SD = 5.99) spent in DTP.

Results from the repeated measures analyses on pre, post, and 6-month follow-up demonstrated that the EDDTP was successful in meeting short-term goals of medical stabilization as measured by BMI (N = 60), (F[1.72, 101.69] = 26.98, p < .001); and normalization of disturbed eating as measured by the EDI-2 DT (N = 41), (F[2, 80] = 12.10, p < .001) and BD (F[1.64, 65.81] = 5.40, p < .011) subscales. The program also demonstrated improved psychological functioning in terms of depression on the CDI (N = 48), (F[1.78, 83.78] = 8.49, p < .002), and in terms of anxiety on the MASC (N = 47), (F[2, 88] = 5.78, p < .005).

Pairwise comparisons (Bonferroni adjusted) were also conducted to examine where the statistical differences lay, and revealed that improvements appear to be maintained (and increased for some measures) at follow-up (see Figures 1–3). Results established that patients’ BMI improved as demonstrated by the significant mean differences between pre and post (p < .001) and pre and 6-month follow-up (p < .001). Patients reported more normalized scores on the DT subscale of the EDI-2, with significant mean differences between pre and post (p < .004) as well as between pre and 6-month follow-up (p < .001). Results revealed that patient body dissatisfaction continued to improve after program completion—the mean difference
from pre to 6-month follow-up was significant ($p < .007$). Patients reported significantly lower depression scores from pre to post ($p < .005$) and from pre to 6-month follow-up ($p < .010$). Anxiety levels improved (although not significantly) during program, and continued to significantly improve after program completion, from pre to 6-month follow-up ($p < .017$). Means and reported ranges on all measures at each time-point are outlined in Table 1.

**FIGURE 1** Youth mean BMI kg/m² across timepoints.

*Note:* BMI: Body mass index (kg/m²).

**FIGURE 2** Youth anxiety and depression across timepoints.

*Note:* --- CDI: Children’s Depression Inventory; — MASC: Multidimensional Anxiety Scale for Children.
Outcomes on the RCI varied by measure (see Tables 2 and 3). The DTP was successful in restoring medical health in the majority of patients, as demonstrated by BMI. From pre to post, 60% moved into the recovered range (meaning their BMIs increased above 19 kg/m²), making a total of 86.9% of patients with a BMI in the healthy range after programming. From post to 6-month follow-up, 9.2% moved into the recovered range, with a total of 64.8% of the patients with healthy BMIs at follow-up. Of importance to note is that over 50% of the patients were classified as having irrelevant change.

**FIGURE 3** Eating Disorder Inventory—Drive for Thinness and Body Dissatisfaction subscale scores across timepoints.

*Note:* – – EDI DT: Eating Disorder Inventory-Drive for Thinness; - - EDI BD: Eating Disorder Inventory-Body Dissatisfaction.

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Medical and Psychological Health Descriptives Across Timepoints: Pre, Post, and at 6-Month Follow-Up</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre day hospital</td>
</tr>
<tr>
<td></td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>BMI kg/m²</td>
<td>18.72 (2.40)</td>
</tr>
<tr>
<td>EDI-DT</td>
<td>16.05 (6.04)</td>
</tr>
<tr>
<td>EDI-BD</td>
<td>19.85 (8.39)</td>
</tr>
<tr>
<td>CDI total</td>
<td>69.85 (16.87)</td>
</tr>
<tr>
<td>CDI total</td>
<td>62.04 (12.22)</td>
</tr>
</tbody>
</table>

*Note.* BMI kg/m²: Body Mass Index; EDI DT: Eating Disorder Inventory—Drive for Thinness Subscale; EDI BD: Eating Disorder Inventory—Body Dissatisfaction Subscale; CDI: Children’s Depression Inventory; MASC: Multi-Dimensional Anxiety Scale for Children.
<table>
<thead>
<tr>
<th>Measures</th>
<th>N</th>
<th>Deteriorated freq (%)</th>
<th>Unchanged freq (%)</th>
<th>Improved freq (%)</th>
<th>Pre to post</th>
<th>Deteriorated freq (%)</th>
<th>Unchanged freq (%)</th>
<th>Improved freq (%)</th>
<th>Pre to post</th>
<th>Pre to post</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI</td>
<td>62</td>
<td>3 (4.6%)</td>
<td>8 (12.3%)</td>
<td>51 (78.5%)</td>
<td>12 (18.4%)</td>
<td>2 (3.1%)</td>
<td>5 (7.7%)</td>
<td>42 (64.6%)</td>
<td>39 (60%)</td>
<td>86.9%</td>
</tr>
<tr>
<td>EDI DT</td>
<td>50</td>
<td>7 (10.8%)</td>
<td>16 (24.6%)</td>
<td>27 (41.5%)</td>
<td>6 (10.7%)</td>
<td>4 (6.2%)</td>
<td>13 (20%)</td>
<td>26 (40%)</td>
<td>13 (20%)</td>
<td>29.5%</td>
</tr>
<tr>
<td>EDI BD</td>
<td>50</td>
<td>11 (16.9%)</td>
<td>22 (35.8%)</td>
<td>17 (26.2%)</td>
<td>15 (23.1%)</td>
<td>4 (6.2%)</td>
<td>15 (23.1%)</td>
<td>16 (24.6%)</td>
<td>11 (16.9%)</td>
<td>42.5%</td>
</tr>
<tr>
<td>CDI</td>
<td>54</td>
<td>9 (13.8%)</td>
<td>15 (25.1%)</td>
<td>30 (46.2%)</td>
<td>9 (13.9%)</td>
<td>7 (10.8%)</td>
<td>12 (18.9%)</td>
<td>26 (40%)</td>
<td>20 (30.8%)</td>
<td>64.2%</td>
</tr>
<tr>
<td>MASC</td>
<td>52</td>
<td>15 (23.1%)</td>
<td>9 (13.8%)</td>
<td>28 (43.1%)</td>
<td>12 (18.5%)</td>
<td>9 (13.8%)</td>
<td>7 (10.8%)</td>
<td>24 (36.9%)</td>
<td>22 (33.8%)</td>
<td>70.6%</td>
</tr>
</tbody>
</table>

*Overall change represents changes for the entire sample regardless of whether the participant's needed to change (i.e. regardless of whether they started above or below the normal cutoff). **Irrelevant change refers to individuals who start off in the normal range on a particular scale, making their positive change irrelevant to the treatment goals of these particular individuals. ***Reliable change represents change of more than 2 standard errors (SE) of measurement for individuals who needed to change (i.e. those whose scores were above the normal cutoff for z-scores and below the cutoff for BMI).
TABLE 3 Reliable Change Index Results and Frequencies for Post Day Program to 6-Month Follow-Up

<table>
<thead>
<tr>
<th>Measures</th>
<th>N</th>
<th>Deteriorated</th>
<th>Unchanged</th>
<th>Improved</th>
<th>Post to 6-months</th>
<th>Deteriorated</th>
<th>Unchanged</th>
<th>Improved</th>
<th>Post to 6-months</th>
<th>Post to 6-months</th>
<th>Post to 6-months</th>
<th>Post to 6-months</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI</td>
<td>61</td>
<td>36 (55.4%)</td>
<td>12 (18.5%)</td>
<td>33 (50.8%)</td>
<td></td>
<td>13 (20%)</td>
<td>8 (12.3%)</td>
<td>6 (9.2%)</td>
<td>64.8%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDI DT</td>
<td>43</td>
<td>9 (13.8%)</td>
<td>16 (24.6%)</td>
<td>14 (20.1%)</td>
<td></td>
<td>5 (7.7%)</td>
<td>12 (18.5%)</td>
<td>13 (13.8%)</td>
<td>46.6%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDI BD</td>
<td>43</td>
<td>6 (9.2%)</td>
<td>16 (24.6%)</td>
<td>15 (23.1%)</td>
<td></td>
<td>2 (3.1%)</td>
<td>10 (15.4%)</td>
<td>8 (12.3%)</td>
<td>82.0%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDI</td>
<td>50</td>
<td>17 (26.2%)</td>
<td>18 (27.7%)</td>
<td>13 (20%)</td>
<td></td>
<td>14 (21.5%)</td>
<td>15 (23.1%)</td>
<td>12 (18.5%)</td>
<td>53.3%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MASC</td>
<td>48</td>
<td>14 (21.5%)</td>
<td>22 (33.8%)</td>
<td>14 (21.5%)</td>
<td></td>
<td>8 (12.3%)</td>
<td>16 (24.6%)</td>
<td>15 (23.1%)</td>
<td>74.8%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

BMI: Body Mass Index; EDI DT: Eating Disorder Inventory-Drive for Thinness Subscale; EDI BD: Eating Disorder Inventory-Body Dissatisfaction Subscale; CDI: Children’s Depression Inventory; MASC: Multi-Dimensional Anxiety Scale for Children. *Overall change represents changes for the entire sample regardless of whether the participant’s needed to change (i.e. regardless of whether they started above or below the normal cutoff). **Irrelevant change refers to individuals who start off in the normal range on a particular scale, making their positive change irrelevant to the treatment goals of these particular individuals. ***Reliable change represents change of more than 2 standard errors (SE) of measurement for individuals who needed to change (i.e. those whose scores were above the normal cutoff for t-scores and below the cutoff for BMI).
from post to 6-months—meaning they did not need to improve because their BMIs were not in the clinically “underweight” range of below 19 kg/m² at post day program. So while only some were in the recovered range, the majority either improved or did not get clinically worse.

Overall, patients were doing well by the end of program in terms of drive for thinness and body dissatisfaction, with 29.5% and 42.5% of patients functioning at a normal/healthy level on these measures at post, respectively. DT scores on the EDI-2 revealed that 20% moved into the recovered range from pre to post, and 13.8% moved into the recovered range from post to 6-month follow-up. Similar results were found for BD, where from pre to post 16.9% moved into in the recovered range, demonstrating clinically significant positive changes in their body dissatisfaction. From post to 6-month follow-up, 12.3% were in the recovered range. Noteworthy, is that at 6-month follow-up, a total of 46.6% were in healthy range for drive for thinness, and 82% were in the healthy functioning range on body dissatisfaction.

Overall, the DTP is shown to be effective in improving feelings of depression, where 30.8% moved into the recovered range, and a total of 64.3% of patients were not deemed clinically depressed post-programming. From post to 6-month follow-up, 18.5% moved into the recovered range. Program benefits were relatively maintained, where over half (53.3%) the patients were deemed clinically healthy when assessed 6-months post-programming. Similar results were found for feelings of anxiety where 30.8% and 23.1% moved into the recovered range from pre to post and at 6-month follow-up respectively. Overall, a total of 70% and 74.8% of patients were not clinically anxious by the end of programming and at 6-month follow-up respectively.

DISCUSSION

The purpose of this article was to provide a descriptive study of a program’s outcomes for youth with moderate to severe eating disorders. The present study contributes to the current literature by demonstrating that a comprehensive family-based DTP can successfully facilitate positive changes in the health of these youth. Findings indicate that, overall, patients in the DTP made clinically meaningful and statistically significant improvements on the three key treatment goals from admission to discharge and at 6-month follow-up.

*Medical stabilization* was achieved as demonstrated by the overall improvement and stabilization in BMI levels across all time points. The change from pre to post, and pre to 6-month follow-up was statistically significant. In terms of clinical significance, by program completion, the majority of patients had improvements in their BMI, and over half had medically recovered (to the point of clinical stabilization). Sixty-three percent
maintained improvements or had irrelevant changes up to 6 months after program completion. Although it may be premature to suggest that this Maudsley based family treatment approach will lead to enduring changes in youth, the overall results on medical stabilization are promising, and they reflect similar diminishments over time to those found in inpatient programs (Collin, Power, Karatzias, Grierson, & Yellowlees, 2010; Steinhausen, Seidel, & Winkler Metzke, 2000).

Normalizing disturbed eating behaviors was achieved both in terms of clinical and statistical significance. Patients reported reduced drives for thinness and body dissatisfaction from pre-programming to 6 months after program completion. RCI analyses further revealed that, overall, many showed some sort of improvement, and that about 1 in 5 recovered on these measures. Follow-up analyses demonstrated continued positive effects of the DTP, with the majority of the sample in a “healthy” range for both DT and BD.

Improving psychological functioning was demonstrated by statistically and clinically significant decreases in depression and anxiety scores throughout treatment and after program completion. Overall close to one-third made some kind of improvement during programming, and about 1 in 5 recovered their psychological functioning into the normal range. At follow up, these improvements continued with the majority of the sample in a “healthy” range for depression and anxiety.

Limitations and Future Directions
The findings describe a successful implementation of the family-based treatment model into intensive treatment, and provide preliminary conclusions on practice-based evidence for the treatment of EDs in youth. While the results from this study are strong and support the success of this treatment, it is noteworthy that 14% were not in a medical recovery recovery range and 54% still struggled with significant drive for thinness at follow-up. These results highlight the need for ongoing therapeutic support for youth and their families in the year following day treatment programs to help youth move forward with their health and to prevent relapse.

This study is restricted by some limitations inherent to clinical research. When dealing with a clinical population and clinical data, there are issues with attrition rates and program half-completers, as well as the use of self-report measures and of missing data due to patient negligence or refusal in completing the measures. The clinical population is also limiting because the program outcomes were evaluated based on assessments of the patients in the program, and therefore there was no control group to use as a comparison. Findings would support a future randomized control trial in order to make comparisons and would improve the generalizability of the
results. Another limitation is that, for the purpose of evaluating outcomes of the overall program, as well as a generally small sample size, patient ED diagnoses were combined and evaluated together, instead of examining progress by diagnostic classification. This is more of a limitation in the ANOVA analysis, as the RCI approach examines change on the individual level and has the advantage of partitioning participants based on whether change is relevant or not. The measurement of medical stabilization is challenging. The use of BMI as a measure of medical stabilization is widely accepted, but the use of blood work, hormone levels etc. could enhance the measurement of medical stabilization. These are more invasive measures that ethically provide limited additional information relative to the invasive procedure required. Future research could look to include further measures of medical stabilization.

Future research could examine a longer follow-up (i.e., 1 or 2 years post program) to determine if changes are maintained beyond 6 months. Further research questions also could be asked, such as: Why do some patients change while others do not? Can we identify/classify patients that are more likely to change? Are there differences in how effective the program seems to be for different ED diagnoses? Findings from this study should not however be undermined by these limitations. Results do characterize some of the realities of research in clinical settings, and highlight key information about the outcomes of a family-based program for youth with eating disorders.

CONCLUSIONS

The findings from the current study suggest that female youth with more severe eating pathology can make not only statistically significant gains from treatment in highly structured treatment settings such as day hospitals, but also clinically meaningful ones. With respect to traditional significance testing, results from the repeated measures analyses and follow-up pairwise comparisons support the effectiveness of the DTP in meeting the short-term treatment goals, and demonstrate maintained benefits after patients leave the secure and stable environment of the day program. The importance of analyzing clinical data according to the meaning of the clinical data is highlighted when examining the difference in the results of the repeated measures and the RCI from post to 6-month follow-up. Based on statistical significance, the repeated measures analyses failed to identify significant changes between post and 6-month follow-up, however, the RCI demonstrated that there are clinically meaningful changes happening. The RCI procedure leads to a more accurate interpretation of outcome assessment by demonstrating that while some changes are occurring, the changes of interest are the clinically significant (recovered) ones. It is of great
importance (to clinicians, researchers, and especially to patients involved in DTPs) to conclude that patients do actually clinically recover in terms of their psychological and/or physical health.

REFERENCES


