Psychodynamic psychotherapy with adolescents and young adults: Outcome in routine practice
Julia Nemirovski Edlund and Gunnar Carlberg
Clin Child Psychol Psychiatry published online 17 October 2014
DOI: 10.1177/1359104514554311

The online version of this article can be found at:
http://ccp.sagepub.com/content/early/2014/10/16/1359104514554311

>> OnlineFirst Version of Record - Oct 17, 2014

What is This?
Psychodynamic psychotherapy with adolescents and young adults: Outcome in routine practice

Julia Nemirovski Edlund¹ and Gunnar Carlberg¹,²
¹The Erica Foundation, Sweden
²Department of Education, Stockholm University, Sweden

Abstract
This naturalistic study examined the outcome of psychodynamic psychotherapy with 218 adolescents and young adults aged 14–24 years. Analysis of variance showed significant improvement of general functioning on Children’s Global Assessment Scale (CGAS) and Global Assessment of Functioning (GAF) and decreased symptom severity on Symptoms Checklist-90 (SCL-90) upon completion of psychotherapy, as well as a clinically significant improvement in a large percentage of cases. Effect sizes were equivalent to those evident in a clinical comparison group and larger than in prior research. The main limitation of this study was the lack of a control group, partially compensated for through the use of comparison groups and high external validity. The study seeks to fill a gap in an important yet overlooked field of research.

Keywords
Psychodynamic psychotherapy, outcome research, child and adolescent psychotherapy, young adults, routine practice

Even though psychodynamic psychotherapy is a well-established method of treatment for adolescents and young adults with psychological disorders, this field of research remains inadequate due to a lack of further systematic and rigorously conducted outcome studies (Midgley & Kennedy, 2011; Reinecke & Shirk, 2007). This study aims to contribute to the field by further investigating the outcome of psychodynamic psychotherapy with a group of adolescents and young adults in routine practice.

Studies performed on groups of adolescent patients have demonstrated an average effect size of 0.7 (Kazdin, 2004; Weisz, Weiss, & Donenberg, 1992). That is a similar effect size to those found in studies performed on adults, implying that findings from studies on adult patients might be generalized to the younger population. Reinecke and Shirk (2007) emphasize that psychopathology in teens often stems from an accumulation of multiple factors and can thus seem complex,
consequently manifesting itself in many different ways. This stands in contrast to the demands posed by the research community to study isolated and well-defined psychiatric disorders and diagnosis, thus making such research not only difficult to generalize to the clinical population but even complicated to carry through.

Several studies of adolescents in therapy have shown that psychodynamic psychotherapy is a more suitable treatment for internalizing than externalizing disorders (Kennedy & Midgley, 2007). In a naturalistic study of Baruch and Fearon (2002), which involved 151 teenagers and young adults, the retention was much higher among those patients who had internalizing symptoms. Tonge, Pullen, Hughes, and Beaufoy (2009) also compared psychodynamic psychotherapy with treatment as usual (TAU) in a longitudinal study on 55 teenagers. The results showed that the largest effect was evident in those who suffered from the most severe symptoms when seeking therapy. These results suggest that adolescents with severe psychiatric disorders might benefit the most from psychodynamic psychotherapy.

There are only a few studies of psychodynamic psychotherapy performed on adolescent patients with specific diagnoses. In the same study by Tonge and co-workers (2009) performed on 72 children and adolescents suffering from depression, 75% were no longer clinically depressed after 9 months of treatment, with none suffering from depression at a 6-month follow-up. In another smaller study by Horn et al. (2005) on 20 children and teenagers with depression or dysthymia, 20% were significantly better after short-term psychotherapy of 25 sessions. No such changes were found in the waiting list control group. A randomized controlled trial (RCT) conducted by Gilboa-Schechtman et al. (2010) compared time-limited dynamic psychotherapy with cognitive behavioral therapy (CBT) for a group of teenagers with posttraumatic stress syndrome (PTSD). PTSD symptoms as well as depression had decreased in both groups, with an overall effect size of 0.87 for the group receiving dynamic psychotherapy. At present, there are only a small number of RCTs of psychodynamic treatment of teenagers with anxiety disorders. Uncontrolled trials, however, suggest that psychodynamic psychotherapy is effective in cases of separation anxiety, phobias, and excessive anxiety (Reinecke & Shirk, 2007).

Young adults are a neglected population within research context (Lindgren, Werbart, & Philips, 2010; Philips, Wennberg, Werbart, & Schubert, 2006; Slavin, 1996). One reason for that might be that many young adults seeking therapy drop out of treatment prematurely, often directly after the assessment stage, generally first 1 to 4 sessions (Slavin, 1996). There are, however, some recent studies of psychotherapy with young adults supporting the effect of psychodynamic treatment. In a study conducted by Falkenström (2010), 80 young adults between the ages of 16 and 23 years completed psychodynamic psychotherapy. Symptoms were self-rated with Symptoms Checklist (SCL-90; Derogatis, 1994), and global functioning was assessed with the Global Assessment of Functioning (GAF) Scale (American Psychiatric Association, 2000). Symptoms decreased significantly after treatment, with an effect size of 1.29. The increase in global functioning was also significant with an overall effect size of 1.11.

Within the Young Adult Psychotherapy Project (YAPP), Lindgren et al. (2010) conducted an outcome study of 92 young adults, who received individual psychodynamic psychotherapy. The results showed a decrease of symptoms after therapy with an overall effect size of 1.33. No correlation was found between outcome and the patients’ gender or age, which corresponded with findings from previous studies (Clarkin & Levy, 2004). Nor did the length of treatment correlate with the outcome. This finding, however, contradicts the findings from a prior meta-analysis of psychodynamic psychotherapy with adults (Leichsenring & Leibing, 2007; Leichsenring & Rabung, 2008). Hence, the findings regarding the association between
length of treatment and outcome are contradictory and should be investigated closely (Kennedy & Midgley, 2007).

The aim of the study

The aim of this study was to examine the outcome of psychodynamic psychotherapy with adolescents and young adults in a naturalistic setting. Research questions were as follows:

1. How can the change in CGAS, GAF, and SCL-90 ratings before and after therapy best be described on group level?
2. How do findings relate to the results from the two following comparison groups: YAPP and SCL-90 population norms?
3. Are the changes described clinically significant?
4. How do therapist- and patient-ratings correlate?

Method

This study is a naturalistic outcome study of psychodynamic psychotherapy with adolescents and young adults, which was conducted as a part of the clinic’s quality assurance work. All patients were informed that data were anonymously gathered for research purposes upon arrival to the clinic. Ethical considerations imposed some limitations on the data collection procedure and data available for research purposes. All therapies in the study were conducted between 2002 and 2009 at an institute providing outpatient psychodynamic psychotherapy for children and adolescents and professional training at university level as well as research. Data analysis was carried out retrospectively using a research database which included information on all patients attending at least one session. Two outcome measures were used: ratings of global functioning were made by the therapists themselves, or at times in cooperation with the senior physician, while the symptoms were self-rated by the patients. Data were collected during assessment (3–4 first sessions) and directly upon termination of therapy.

Participants

The sample included 218 teenagers and young adults of 14–24 years of ages who underwent assessment (three to four sessions) and consequently were offered psychotherapy. Patients had, in most cases, been allocated to the next available therapist. An additional inclusion criterion was a minimum of attended sessions, which was set to six, thereby excluding all patients attending solely the assessment phase. The majority of the patients were females (76%), and the average age was 19.17 years ($s = 2.45$). Table 1 shows the participants’ age and gender.

Table 1. Age group and gender.

<table>
<thead>
<tr>
<th>Age group</th>
<th>Women ($n=166$)</th>
<th>Men ($n=52$)</th>
<th>Total ($N=218$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>$%$</td>
<td>$n$</td>
</tr>
<tr>
<td>14–19</td>
<td>94</td>
<td>43.1</td>
<td>33</td>
</tr>
<tr>
<td>20–24</td>
<td>72</td>
<td>33.0</td>
<td>19</td>
</tr>
</tbody>
</table>
Diagnoses

During the assessment process, the therapist assigned each patient with a diagnosis according to the axis-I of the Diagnostic and Statistical Manual of Mental Disorders (4th ed.; DSM-IV; American Psychiatric Association, 2000). Diagnoses were subsequently divided into seven categories. Those categories were then entered into the clinic’s research database, which this study builds upon. Therefore, the patients’ individual specific diagnoses were not available at the time of the study. A third of the patients were assigned with a mood disorder as a primary diagnosis, and a third with other diagnoses ($n=69$). Nearly a third were diagnosed with anxiety disorders ($n=59$). A total of 12 participants did not fulfill the criteria for any particular psychiatric diagnosis. However, they presented with a severe enough degree of complaints to motivate psychotherapy. Data about diagnosis were missing in four cases. One-fourth of the participants (27%) were assigned with more than one diagnosis during assessment, also described as comorbidity, which was accordingly evident in 59 cases. Only axis-I diagnoses were assigned due to the participants’ young age as well as the clinic’s general policy requiring only axis-I diagnoses.

Psychosocial and environmental problems

During assessment, the therapists were able to identify psychosocial and environmental factors that contributed to the patients’ psychological disorders according to the fourth axis of the DSM-IV (American Psychiatric Association, 2000). Nearly half of the participants reported problems with primary support group, such as conflicts between adults in the family ($n=104$). Problems related to the social environment were identified in just over a third of the cases ($n=77$). Other common problems were substance abuse in the family ($n=45$), physical or emotional health problems in the family ($n=31$), or death of a family member ($n=29$).

Measures

Symptoms Checklist-90. This is a self-rating questionnaire with 90 statements regarding symptoms, ranging on a scale from 0 (not bothered at all) to 4 (bothered a great deal) (Derogatis, 1994). Global Severity Index (GSI), which is the average of all 90 items, was used as an outcome measure. The instrument has shown good reliability with Cronbach’s $\alpha=.97$ (Fridell, Cesarec, Johansson, & Thorsen, 2002).

Children’s Global Assessment Scale. This instrument was used to rate the general functioning of adolescents under the age of 20 years on a scale from 1 to 100 divided into 10-point intervals, each with a descriptive title followed by concrete examples (Shaffer et al., 1983). The translation of the scale has been validated and is widely used both within clinical work as well as in research (Lundh, Kowalski, Sundberg, Gumpert, & Landén, 2010).

GAF. The GAF scale constitutes the fifth axis of the DSM-IV classification system and aims to rate social, occupational, and psychological functioning of the patient (American Psychiatric Association, 2000). In this study, GAF ratings were made for patients aged 20 years or older by the therapists themselves. The scale’s reliability has proven to be very high in research context, whereas the reliability in clinical studies is clearly lower (Ramirez, Ekselius, & Ramklint, 2008). In spite of varying results, the scale is regarded reliable and is widely used in psychotherapy assessment.
Therapies

All therapies were conducted either by psychodynamically oriented psychotherapists who were specialized in psychodynamic psychotherapy or by psychotherapy students with prior psychotherapy experience operating under the supervision of specialists and trained supervisors. Blagys and Hilsenroth (2000) identified seven unique characteristics of psychodynamic psychotherapy in a review comparing empirical studies of psychodynamic therapy with CBT:

1. Focus on affect and expression of emotion;
2. Exploration of attempts to avoid distressing thoughts and feelings;
3. Identification of recurring themes and patterns;
4. Discussion of past experience (developmental focus);
5. Focus on interpersonal relations;
6. Focus on therapy relationship;
7. Exploration of fantasy life.

The length of the treatment varied between 6 and 348 sessions of 45 minutes with 43 sessions on average (standard deviation \(SD = 50\)). The median for the total number of sessions was 26. The majority of patients \(n = 186\) attended treatment once a week, 13 patients attended twice weekly, and 19 attended with varying frequency. Therapies were divided into three groups: short-term (6–24 sessions), middle-length (25–50 sessions), and long-term (>50 sessions). Table 2 shows therapy frequencies respectively.

Comparison groups

Results were compared with two groups:

1. Young adults aged 18–25 years who participated in the YAPP within the Stockholm Institute of Psychotherapy. The group comprised 134 young adults who received psychodynamic individual or group treatment and of which 97 (73%) were women (Philips et al., 2006).
2. SCL-90 national norms based on a representative sample of 2776 adolescents (2305 girls and 471 boys) aged 17–19 years and 386 young adults (337 women) aged 20–25 years (Fridell et al., 2002).

Statistical analysis

All statistical analysis was conducted in IBM SPSS Statistics 20.0. The outcome of psychotherapy at group level was analyzed with mixed-design analysis of variance (ANOVA) (General Linear Model–Repeated measures) with one within fixed-factor (time). The same method was used to

<table>
<thead>
<tr>
<th>Table 2. Treatment length.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-term</td>
</tr>
<tr>
<td>6–24 sessions</td>
</tr>
<tr>
<td>(n)</td>
</tr>
<tr>
<td>(%)</td>
</tr>
<tr>
<td>105</td>
</tr>
<tr>
<td>48.2</td>
</tr>
</tbody>
</table>
interpret interaction effects of time and each one of the grouping variables: age group (adolescents 14–19 years/young adults 20–24 years), gender, length of therapy (short-term, middle-length, long), and comorbidity (two groups: one or none diagnosis/at least two diagnoses). In order to minimize the effect of patient dropout on outcome, an last observation carried forward (LOCF) imputation method was adopted. Baseline observations were subsequently carried forward to the posttreatment observations for those patients missing posttreatment data. Bonferroni correction of the alpha was performed for the repeated comparisons across the outcome variables. Alpha level was consequently set to .025 for each test. One-way ANOVA was then conducted in order to further explore the interaction of time and length of therapy. In addition, a set of paired sample t-tests were performed to explore change over time for each one of the major diagnosis groups (anxiety disorders, mood disorders, and other diagnoses) and for each outcome measure. In those analyses, alpha level was set to .008. Nonetheless, due to the risk of spurious significances, results should be interpreted with caution.

Normative comparison of posttreatment scores on the SCL-90 as well as comparison to the clinical comparison group was conducted using null-hypothesis testing as discussed by Kendall and Grove (1988). Consequently, both pre- and posttreatment scores on the SCL-90 and the GAF scale were compared to the results from the clinical comparison group reported by Philips et al. (2006) within the YAPP project.

Cohen’s d (effect size) was obtained to describe the amount of change across the two time-points using the following formula: \(\left(\frac{\text{Mean}_{\text{before}} - \text{Mean}_{\text{after}}}{SD_{\text{pretreatment}}}\right)\). An effect size of 0.8 was considered large, 0.5 medium, and 0.2 small (Cohen, 1988).

Correlations between therapists’ and patients’ ratings, before and after therapy, respectively, were calculated with the Pearson product–moment correlation coefficient.

Reliable and clinically significant improvement according to Jacobson and Truax (1991) was used as a measure of individual change over the two time-points. Patients who during the study crossed the limit from a clinical into a normal population, having changed to a degree not attributable to measurement error, were stated as achieving clinically significant change. Reliable Change Index (RCI) for SCL-90 was calculated for each individual patient using reliability coefficients reported by Schmitz, Hartkamp, and Franke (2000). The following cutoff limits suggested by Falkenström (2010) were used: \(c=0.95\) for women 16–19 years of age, \(c=0.84\) for women 20–25 years of age, \(c=0.66\) for men 16–19 years of age, and \(c=0.71\) for men 20–25 years of age. For Children’s Global Assessment Scale (CGAS) and GAF scales RCI was of 10 points. Cutoff limits were set to 70 based on prior literature (Schorre & Vandvik, 2004).

### Attrition

Out of the 218 patients who fulfilled the inclusion criteria regarding age (14–24 years) and number of sessions (minimum of six including assessment), 49 (22.5%) submitted SCL-90 questionnaires with more than 20% missing items during assessment and were consequently, as advised by Fridell et al. (2002), excluded from the analysis. Thus, complete initial SCL-90 ratings were obtained from 169 patients. Initial GAF and CGAS ratings of eight patients (four of each respectively) were missing. Therefore, the final analysis included 114 CGAS ratings and 96 GAF scale ratings (Figure 1).

Patients with missing SCL-90 data pre-therapy were compared with patients who submitted complete SCL-90 questionnaires, on the following seven variables: gender, age, length of therapy, number of diagnoses, comorbidity, number of psychosocial and environmental problems, and global functioning. No statistically significant differences were found between the groups.

An additional analysis was then conducted on the eight patients with missing C-GAS and GAF ratings. Consequently, it had been established that patient characteristics (age, gender, diagnosis, psychosocial and environmental problems, and comorbidity) of those eight participants did not
differ from those of the group of patients included in the study, implying that their attrition did not affect outcome significantly.

**Results**

*Change at group level*

Data analysis has shown a significant main effect of time (before/after psychotherapy) for all three measurements: SCL-90, CGAS, and GAF. Wilks’ Lambda = 0.74, F(1, 160) = 57.21, p < .001 for GSI (SCL-90) as a dependent variable, Wilks’ Lambda = 0.62, F(1, 108) = 66.57, p < .001 for CGAS, and Wilks’ Lambda = 0.41, F(1, 89) = 130.80, p < .001 for GAF. Both the GAF and CGAS scales generated large effect sizes, $d = 2.02$ for GAF and $d = 1.54$ for CGAS.

The average GAF scale rating for the whole group before therapy was 55.09, which is within the category labeled “Moderate symptoms OR moderate difficulty in social, occupational, or school functioning” (American Psychiatric Association, 2000). The group’s mean on the GAF scale was significantly lower pretherapy ($t_{190} = 2.3, p < .05$) than that reported by Philips et al. (2006). After therapy, the group’s average GAF score had increased to 68.54, which is within the interval described as “Some mild symptoms OR some difficulty in social, occupational, or school functioning, but generally functioning pretty well, has some meaningful interpersonal relationships.” A similar increase in general functioning could be observed on the CGAS scale where the pretherapy average was 55.86 which is described as

Variable functioning with sporadic difficulties or symptoms in several but not all social areas; disturbance would be apparent to those who encounter the child in a dysfunctional setting or time but not to those who see the child in other settings. (Shaffer et al., 1983)
The average CGAS score post therapy had increased to 67.36, within the interval described as some difficulty in a single area but generally functioning pretty well; has some meaningful interpersonal relationships; most people who do not know the child well would not consider him/her deviant but those who do know him/her well might express concern. (Shaffer et al., 1983)

The patients’ mean self-rated symptoms decreased significantly after therapy, with an effect size of 0.76. Average GSI score (SCL-90) prior to therapy was 1.29, which was much higher than the mean of the normal population as described by Fridell et al. (2002) indicating severe problems. After therapy, the mean GSI score had decreased to 0.81. Similar figures were reported within the YAPP project (Philips et al., 2006), where the average GSI before therapy was 1.3, and 0.8 after. Independent t-tests comparing posttherapy scores on SCL-90 and GAF scale to those reported within the YAPP project showed no significant differences. Results are shown in Table 3.

Furthermore, test of between-subjects effects revealed a main effect of comorbidity for two of the dependent variables. For SCL-90: $F(1, 160) = 14.45, p < .001$, and for CGAS: $F(1, 108) = 8.02, p = .006$. These findings are not surprising given that the number of diagnoses often go hand in hand with the general level of functioning and symptom severity, consequently strengthening the validity of the measuring instruments. No main effects of either gender or age were found. Nor were any interaction effects discovered between time and any of the following grouping variables: gender, age, and comorbidity.

A statistically significant interaction effect was found between time and length of treatment for both dependent variables CGAS and GAF. Wilks’ Lambda $= 0.93, F(2, 108) = 4.11, p = .019$ for CGAS and Wilks’ Lambda $= 0.84, F(2, 89) = 8.37, p < .001$ for GAF. However, no such interaction was discovered when measured with SCL-90. Further exploration of those effects with a one-way ANOVA revealed significant differences in both GAF and CGAS scores between those receiving short-term psychotherapy (6–24 sessions) and long-term therapy (>50 sessions) suggesting greater change in those receiving longer treatments.

An additional analysis of the data was made for each major diagnosis category: anxiety disorders, mood disorders, and other diagnoses. The results of the paired sample t-tests yielded similar results to the whole-group analysis, with significant increase in CGAS and GAF scores after therapy for all three groups as well as a significant decrease in GSI scores and large effect sizes throughout. Table 4 shows change in GSI, CGAS and GAF scores before therapy and at termination for each diagnosis group separately.

The patients’ GSI scores were compared to the national norms as presented by Fridell et al. (2002). Table 5 shows GSI scores of men and women separately, divided into two age groups in Table 3.

### Table 3. Changes in mean level of functioning (CGAS/GAF) and self-rated symptoms (SCL-90) from beginning to end of therapy (SD in parentheses).

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Before therapy</th>
<th>After therapy</th>
<th>Wilks’ Lambda</th>
<th>p</th>
<th>Effect size</th>
<th>Posttherapy means (Philips, Wennberg, Werbart, &amp; Schubert, 2006)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCL-90</td>
<td>169</td>
<td>1.29 (0.63)</td>
<td>0.81 (0.55)</td>
<td>0.74</td>
<td>.001</td>
<td>0.76</td>
<td>0.8 (0.6)</td>
</tr>
<tr>
<td>GAF</td>
<td>96</td>
<td>55.09 (6.66)</td>
<td>68.54 (12.02)</td>
<td>0.41</td>
<td>.001</td>
<td>2.02</td>
<td>67.3 (11.4)</td>
</tr>
<tr>
<td>CGAS</td>
<td>114</td>
<td>55.86 (7.49)</td>
<td>67.36 (11.25)</td>
<td>0.62</td>
<td>.001</td>
<td>1.54</td>
<td>N/A</td>
</tr>
</tbody>
</table>

CGAS: Children’s Global Assessment Scale; GAF: Global Assessment of Functioning; SCL: Symptoms Checklist; SD: standard deviation.
addition to the scores of the normal population for each group. All four groups reported much higher than normal severity of symptoms before beginning therapy. Upon completion of treatment, the GSI of 73 women aged 14–19 years decreased markedly to the level of same aged women in the normal population \((t_{2376} = 0.5, \text{n.s.})\). The other three groups’ GSI scores clearly decreased as well, though remaining somewhat above the scores of the normal population.

**Individual change**

Clinically significant change describes a transition between clinical and normal population as well as the amount of change obtained (Jacobson & Truax, 1991). Out of 169 patients with complete SCL-90 data, 113 (67%) achieved reliable change, 41 (24%) were unchanged, and 15 (9%) deteriorated reliably. Prior to therapy, 123 (73%) patients had symptoms’ severity above the cutoff level and were therefore classified as belonging to a clinical population range. Upon completion of therapy, half of the patients in this group have crossed the cutoff limit into the normal population range and changed reliably, thus indicating clinically significant change. A small group of three patients, who initially belonged to the normal population, with GSI scores under the SCL-90 cutoff limits, deteriorated reliably posttreatment, while crossing the limit from normal into the clinical sphere.

For the CGAS and GAF scales, a cutoff limit of 70 was set. While 69 (61%) patients improved reliably on the CGAS, 45 (39%) remained unchanged. Similar numbers were obtained on the GAF scale, with 63 (66%) reliably improved and 33 (34%) unchanged. None of the patients showed reliable deterioration on the therapist-rated scales. During assessment, a majority of patients (96%) were assigned a CGAS score below 70, which is within the clinical population. Nearly half of this group (43%) improved significantly and no longer belonged to the clinical population post therapy. A similar change was evident for GAF, where 51% improved significantly, crossing the cutoff limit into the normal population and changing to a larger extent than could be attributed to measurement error. None of the patients’ general functioning according to CGAS or GAF deteriorated clinically.

**Table 4.** SCL-90, CGAS, and GAF means and standard deviations before and after therapy divided into three diagnosis categories: anxiety disorders, mood disorders, and other diagnoses.

<table>
<thead>
<tr>
<th>Scale</th>
<th>n</th>
<th>Before therapy</th>
<th>After therapy</th>
<th>t</th>
<th>p</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCL-90</td>
<td>53</td>
<td>1.26</td>
<td>0.77</td>
<td>5.83</td>
<td>.001</td>
<td>0.76</td>
</tr>
<tr>
<td>GAF</td>
<td>30</td>
<td>55.07</td>
<td>71.87</td>
<td>8.77</td>
<td>.001</td>
<td>2.98</td>
</tr>
<tr>
<td>CGAS</td>
<td>28</td>
<td>54.89</td>
<td>67.36</td>
<td>8.66</td>
<td>.001</td>
<td>1.47</td>
</tr>
<tr>
<td>Mood disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCL-90</td>
<td>51</td>
<td>1.41</td>
<td>0.90</td>
<td>6.48</td>
<td>.001</td>
<td>0.82</td>
</tr>
<tr>
<td>GAF</td>
<td>33</td>
<td>54.09</td>
<td>65.00</td>
<td>6.06</td>
<td>.001</td>
<td>1.72</td>
</tr>
<tr>
<td>CGAS</td>
<td>36</td>
<td>53.81</td>
<td>66.00</td>
<td>6.68</td>
<td>.001</td>
<td>1.80</td>
</tr>
<tr>
<td>Other diagnoses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCL-90</td>
<td>51</td>
<td>1.15</td>
<td>0.73</td>
<td>5.60</td>
<td>.001</td>
<td>0.73</td>
</tr>
<tr>
<td>GAF</td>
<td>27</td>
<td>56.30</td>
<td>68.85</td>
<td>6.95</td>
<td>.001</td>
<td>1.56</td>
</tr>
<tr>
<td>CGAS</td>
<td>39</td>
<td>59.00</td>
<td>69.54</td>
<td>7.83</td>
<td>.001</td>
<td>1.58</td>
</tr>
</tbody>
</table>

SCL: Symptoms Checklist; CGAS: Children’s Global Assessment Scale; GAF: Global Assessment of Functioning.
Table 6. Patients achieving reliable and clinically significant change upon completion of treatment.

<table>
<thead>
<tr>
<th>Amount of change</th>
<th>CGAS</th>
<th>GAF</th>
<th>SCL-90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reliable improvement</td>
<td>61% (69/114)</td>
<td>66% (63/96)</td>
<td>67% (113/169)</td>
</tr>
<tr>
<td>No change</td>
<td>39% (45/114)</td>
<td>34% (33/96)</td>
<td>24% (41/169)</td>
</tr>
<tr>
<td>Reliable deterioration</td>
<td>0% (0/114)</td>
<td>0% (0/96)</td>
<td>9% (15/169)</td>
</tr>
<tr>
<td>Clinically significant</td>
<td>43% (47/109)</td>
<td>51% (48/94)</td>
<td>51% (63/123)</td>
</tr>
<tr>
<td>Improvement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinically significant</td>
<td>0% (0/5)</td>
<td>0% (0/2)</td>
<td>7% (3/46)</td>
</tr>
<tr>
<td>deterioration</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>114</td>
<td>96</td>
<td>169</td>
</tr>
</tbody>
</table>

CGAS: Children’s Global Assessment Scale; GAF: Global Assessment of Functioning; SCL: Symptoms Checklist.

Relationship between therapists’ and patients’ ratings

Therapists’ ratings were significantly correlated with the patients’ self-ratings both at assessment and termination. There was a moderate negative correlation between therapists’ CGAS ratings and patients’ GSI scores prior to therapy ($r = -0.36; p < .001$). GAF scores, assigned to the patients before therapy, correlated weakly with the patients’ GSI ratings ($r = -0.26; p = .028$). Upon completion of therapy, the correlation between CGAS and GSI scores remained at the same intermediate level ($r = -0.34; p = .002$). However, the correlation between GAF ratings and GSI increased ($r = -0.48; p < .001$). Philips et al. (2006) reported similar although somewhat bigger variation in their study, with a weak correlation, $r = -0.29$, prior to therapy and large correlation between therapists’ and patients’ ratings, measuring up to $r = 0.60$, upon completion.

Discussion

This study aimed to examine the outcome of psychodynamic psychotherapy with adolescents and young adults within a clinical setting. Results showed a significant improvement of general functioning, measured with the CGAS and GAF scales, as well as a decreased symptom severity according to self-rated SCL-90 upon completion of treatment. In comparison to national SCL-90 norms (Fridell et al., 2002), GSI scores of 73 women aged 16–19 years decreased considerably to the same level as that of same aged women in the normal population. Furthermore, the patients’ improvements on both measurements were comparable to the results reported in the YAPP project (Philips et al., 2006). The current study’s effect sizes on all three measurements were considerably larger than those shown in previous meta-analyses and reviews of psychodynamic psychotherapies.
with adolescents, which measured up to 0.7 on average (Kazdin, 2004; Weisz et al., 1992). The same trend was observed when comparing the results to single studies of psychodynamic treatment with young people (Gilboa-Schechtman et al., 2010; Kronmüller et al., 2005). In accordance to prior research, neither age nor gender correlated with therapy outcome (Clarkin & Levy, 2004; Leichsenring & Rabung, 2008; Lindgren et al., 2010).

An interaction effect of time and length of therapy was found for both measurements CGAS and GAF, suggesting greater improvement in general functioning of patients receiving long-term treatment (>50 sessions) than those receiving short-term treatment (6–24 sessions). However, those results should be interpreted with caution. Dropout from treatment is known to occur at a higher rate in the initial phases of therapy (Slavin, 1996). Thus, dropouts might have led to lesser degree of change in the short-term treatment group, although some research works suggest that premature termination of treatment does not necessarily lead to a less favorable outcome (Bobart & Greaves Wade, 2013). Similar results were shown in a previous review of Leichsenring and Leibing (2007) stating longer and more frequent psychotherapy to be more effective than shorter treatment. In yet another meta-analysis, Leichsenring and Rabung (2008) reported a significant interaction between outcome and number of sessions in long-term psychotherapy. However, in the same study, outcome did not correlate with treatment duration measured in weeks. Other studies, nevertheless, have reported conflicting results, at times finding no connection between outcome and therapy length (Kennedy & Midgley, 2007). Contradictory results suggest complicated interaction patterns between outcome and length of treatment and should therefore be investigated closely.

Analysis of individual change by means of clinical significance revealed a clinically significant improvement in both symptoms’ severity and general functioning of patients. Proportions of reliably and clinically improved patients were comparable to those recently reported by Falkenström (2010) as well as other previous studies (Baruch & Fearon, 2002; Kronmüller et al., 2005; Target & Fonagy, 1994; Tonge et al., 2009). However, percentages of clinically significant improvement reported in prior literature differ greatly, mainly due to the variation in the measurement methods applied, as well as the limited usage of clinical significance in prior research in general, hence making comparison difficult.

Nevertheless, not all patients benefited from treatment. A small group of three patients with symptom severity levels within the normal population range prior to therapy deteriorated, thus crossing the cutoff limit into the clinical population upon completion of treatment. However, due to the demographic heterogeneity of this group coupled with a variation in diagnoses assigned to the patients during assessment, any valid conclusions regarding those results could not be drawn. No such deterioration was evident in the general functioning of the same patients according to the CGAS or GAF scales. One plausible explanation would be that adolescents and young adults, as part of their age appropriate development, often present with psychological symptoms to a higher extent than that of adults. Using symptoms’ load as an outcome measure for this age group therefore runs the risk of generating misleading results. Although not discussed in this study, harmful effects of psychotherapeutic treatment is an important yet overlooked issue in psychotherapy research, which deserves a closer investigation (Dimidjian & Hollon, 2010).

The study’s main strength is its naturalistic design, which increases the external validity of the reported findings. The wide inclusion criteria in combination with the heterogeneity of the investigated group of subjects bring the study closer to clinical reality, thus allowing greater generalizability of findings than that of strictly randomized controlled studies. In addition, this study aims to address a population of adolescents and young adults, otherwise often overlooked in research context. The study not only replicates recent studies carried through on this group but also broadens the evidence base regarding those particular age groups.
An important limitation of this study is the lack of a follow-up, which was not possible to carry through due to the retrospective nature of the design. In open-ended psychodynamic psychotherapies, therapist and patient often decide together when to terminate the treatment. It is most likely that therapies are terminated at a time when the patient reports fewer symptoms and better functioning, hence leading to more favorable outcome scores on the different measurements and larger effect sizes. However, although there is no evidence to support the applicability of these findings to this study, previous studies have repeatedly shown that the benefits of psychodynamic psychotherapy remain and increase significantly several months after termination (Shedler, 2010). The lack of a control group could be partially compensated for by the use of an equivalent clinical comparison group as well as normative comparison as discussed by Kendall and Grove (1988). Nevertheless, those methods do not alleviate the potential effects of regression to the mean or natural course of recovery.

All data analysis in this study was conducted using a large research database which was constructed for the clinic’s general quality assurance purposes regardless of this study’s specific aim and research questions. Therefore, and due to ethical considerations, any further data regarding therapists, dropouts, and assessment were not available, which imposed significant limitations. Therapist’s effects and further dropout analysis were for that reason not possible to carry through. An additional major limitation is that the treating therapists themselves conducted assessment rather than blind raters. Clinician-based ratings are known to report greater change than that obtained through client-based ratings (Cuijpers, Li, Hofmann, & Andersson, 2010). However, large effect sizes were obtained in spite of the use of client-based ratings on the SCL-90, which in turn significantly correlated with the clinician-based ratings on CGAS and GAF both prior to and post treatment. Furthermore, large effect sizes were obtained throughout despite the highly conservative imputation method (LOCF) and the inclusion of subclinical patients without psychiatric diagnoses, factors that are otherwise known to substantially reduce effect size (Barkham, Stiles, Connell, & Mellor-Clark, 2012).

In conclusion, countless reports published during past years have highlighted the increasing deterioration of mental health in young people (Reading, 2007). Nonetheless, adolescents and young adults are often overlooked in clinical research, which largely focuses on the adult population (Midgley & Kennedy, 2011; Reinecke & Shirk, 2007). Furthermore, during recent years, randomized controlled studies have been the gold standard for psychotherapy research, making findings harder to generalize to the clinical population. As pointed out by Weisz, Ugueto, Cheron, and Herren (2013),

. . . one possible reason evidence based psychotherapies may not fare so well with referred youths who have serious problems, in comparison to usual care in practice settings, may be that so little of the treatment research in our field has been done with clinically referred youngsters treated in everyday practice settings. (p. 275)

This study seeks to fill those two gaps in a young yet important research area of psychodynamic psychotherapy with adolescents and young adults in routine practice, thus contributing to the evidence base for psychodynamic psychotherapy as an effective treatment method for young people with various psychological disorders.

Acknowledgements
The authors wish to thank the clinical team at the Erica Foundation, Stockholm.

Funding
This study was supported by “Gålöstiftelsen.”
Note
1. Diagnoses were divided into seven categories, each category given below is followed by several examples:

1) **Anxiety disorders.** Panic disorder, obsessive-compulsive disorder, phobia, posttraumatic stress syndrome (PTSD), or generalized anxiety disorder.

2) **Attention deficit disorders.** Conduct disorder not otherwise specified (NOS), attention-deficit/hyperactivity disorder, oppositional defiant disorder, or conduct disorder.

3) **Mood disorders.** Major depressive disorder, dysthymia, depression NOS, or bipolar disorder.

4) **Other disorders of infancy, childhood, or adolescence.** Separation anxiety disorder, reactive attachment disorder, or disorder of infancy, childhood, or adolescence NOS.

5) **Other diagnoses.** Identity problem, eating disorders, sleep disorders, or somatic problems.

6) **Pervasive developmental disorders.** Asperger’s disorder or autistic disorder.

7) **No diagnoses:** No diagnoses established according to the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; DSM-IV) criteria.

References


**Author biographies**


**Gunnar Carlberg,** PhD, Professor at the Department of Education, Stockholm University and former Director of the Erica Foundation in Stockholm. He is a licensed psychologist and psychotherapist with various publications within psychodynamic developmental psychology and child psychotherapy research.