Effectiveness of family therapy: a meta-analysis

Evelyn Markus*, Alfred Lange* and Thomas F. Pettigrew

The effectiveness of family therapy is assessed in a meta-analysis of 19 studies. At post-treatment, the average patient with family therapy is better off than 76% of the patients with an alternative treatment, a minimal treatment or no treatment. This effect is comparable to that assessed of other psychotherapies at post-treatment. Our data suggest, that the effect of family therapy increases during the first year after treatment, but that the effect may diminish sharply 18 months after the end of therapy. However, more studies with extended follow-ups will be needed before firm conclusions about the long-term effectiveness of family therapy can be advanced.

Along with the growth of family therapy in mental health service, data on its effectiveness have also increased. The literature on these studies has been reviewed several times in order to integrate the different results and to estimate the overall effectiveness of family therapy (Gurman et al., 1986). In 1987, the first quantitative literature review or 'meta-analysis' on family therapy effectiveness was published (Hazelrigg et al., 1987). During the last decade, meta-analysis enjoyed a growing popularity as a review method in several areas of psychology as well as in other sciences. It provides a highly quantitative method for integrating research results and yields conclusions that are also more quantitative, and hence more precise and less arbitrary than the conclusions of most traditional narrative review methods (Rosenthal, 1984). Meta-analysis rests on the basic assumption that it is in the same spirit to pool the outcomes of a group of studies as to pool the outcomes of a group of individual subjects in a sample. Instead of studying a sample of individual subjects, one studies a sample of effect studies, in which each study outcome is seen as

* Evelyn Markus, M.A., is a fellow-worker in Clinical Psychology of the University of Amsterdam; Alfred Lange, Ph.D., is an Associate Professor of Clinical Psychology at the University of Amsterdam; Thomas F. Pettigrew, Ph.D., is Professor of Social Psychology at the University of Amsterdam and the University of California, Santa Cruz.

Please address all correspondence to: Alfred Lange, Department of Clinical Psychology, Weesperplein 8, 1018XA Amsterdam, The Netherlands.

0163-4445/90/030205 $03.00/0 © 1990 The Association for Family Therapy
if it were one observation in a sample. The results are combined into an
overall effect size and can be analysed with the same type of statistical
methods that are usually applied in individual effect studies. Thus, the
literature review can be as quantitative as the individual effect studies. It
even has the advantage that each observation in fact is based on a whole
sample of observations.

The meta-analysis of Hazelrigg et al. included 20 studies with controls.
At post-treatment, family therapy turned out to have a so-called ‘moderate’ effect size of \( d = 0.46 \); this means that the average patient with
family therapy scored 0.46 standard deviations better than the average
patient in the control groups. This effect seemed to decrease somewhat
over time \( (r = -0.30) \).

We present in this article another meta-analysis of family therapy
outcome studies, conducted a year after that of Hazelrigg et al. Ten of the
studies included in our meta-analysis were also used by Hazelrigg et al.,
and the other studies are new. Altogether 19 studies are included. In this
replication of Hazelrigg et al. (1987) we estimate the overall effectiveness
of family therapy and compare this to other therapies.

**Method**

**Sample of studies**

We searched for family therapy outcome studies in *Psychological Abstracts*
and *Dissertation Abstracts*, both by manual and computer search. In
addition, the reference lists of previous reviews of family therapy
outcome studies were used to locate articles and unpublished material.
Unpublished manuscripts are important in a meta-analysis in order to
prevent ‘publication bias’, as studies with significant results may be more
likely to be accepted for publication. The search involved the period
from January 1967 to May 1987, and uncovered 197 articles. Only
studies about *conjoint* family therapy were included in the meta-analysis.
Studies of family therapy in groups or studies of marital therapy with no
children involved were not included. As a final requirement, study
reports had to contain results and statistical analyses that were necessary
for meta-analytic computations. This meant that the retained studies
contained either a measure of effect size (i.e. \( r \) or \( d \)), the results of tests of
significance (i.e. \( t \) or \( F \)), or a sufficient amount of raw data for effect sizes
to be calculated. Studies without control group could also be included.

Only 17 papers, all published in English Language journals, met these
criteria and could be included in the meta-analysis. One article
contained the reports of three separate studies (Barton et al., 1985), so
that 19 independent studies could be included in the meta-analysis. Table 1a through 1g provides an overview. Many of the other publications that were located by our search could not be included in our meta-analysis, because they lacked the required minimal statistical information (i.e., n- and p- values of the hypothesis testing), or because they were published in (volumes of) journals that were not available in any of the Dutch libraries. Three studies that were included in the meta-analysis of Hazelrigg et al. (1987) were left out of our analysis because they did not meet our selection criteria [parent-group training, multiple family therapy (Stover and Guerny, 1967; Stuart et al., 1976; Hardcastle, 1977)]. Seven other studies which were included by Hazelrigg et al., were not available in Dutch libraries. No unpublished material was located.

**Measuring the effect**

First the results of each study were transformed to one common measure of effect size, Cohen's $d$ (Cohen, 1977). In this measure, the difference between the mean score of the experimental group and the mean score of the control group was expressed in standard units, 

$$d = \frac{(X_c - X_e)}{s},$$

where $X_c$ = the mean post score of the experimental group, $X_e$ = the mean post score of the control group and $s$ = the pooled standard deviation of both groups. $D$ becomes negative if the post scores are less favourable for the experimental group than for the control group. The group receiving family therapy was labelled as the experimental group. If a design contains more than one control group, the scores of these control groups were pooled. Only post scores were included in the analysis. In studies without any kind of control group, however, measurements made before were compared with those measurements made after.

If a study measured the effect in various ways, we did not always include all measurements in the analysis. Instead, for each dependent variable we used the data from only one measurement instrument. The criteria for selecting these instruments are described in detail by Markus (1988).

When Cohen's $d < 0.20$, the effect was interpreted as 'small'; when $0.20 < d < 0.80$, the effect was interpreted as 'moderate'; and an effect of $d > 0.80$ was interpreted as 'big' (Rosenthal, 1984).

The formula $d = \frac{(X_c - X_e)}{s}$ could only be directly applied if means and standard deviations of the post scores were explicitly reported. In most articles, however, this information was omitted. Results were usually presented in values of $t$, $F$, chi-square or $P$. The results of these significance
E. Markus, A. Lange and T. F. Pettigrew

Tests could be transformed to a value of Cohen's $d$ by the formula given by Rosenthal (1984).

Tables 1a through 1g provide an overview of the 19 studies and their separate outcomes expressed in Cohen’s $d$. For each study one $d$-value was computed regarding post-treatment (0–2 months after the end of therapy) and one $d$-value regarding the follow-up. Before computing the combined effect size, the $d$-values of Table 1 were corrected for sample size.

The results of all separate studies were then combined by taking their average and testing with three statistical tests: Stouffer's Combined Test of significance (Stouffer et al., 1949), a chi-square test for homogeneity and a ‘fail-safe’ test for publication bias (Rosenthal, 1984). Therapy outcome was computed separately for post-treatment and for follow-up.

**Results**

Tables 1a and b provide an overview of the separate outcomes in studies where family therapy is compared with a no-treatment control condition, using behaviour ratings and recidivism respectively. Table 1c, d and e contain studies in which family therapy is compared with a minimal treatment control group in which patients were engaged in attention placebo sessions or put on a waiting list. These three tables list studies that are measuring family interaction, behaviour ratings and recidivism respectively. In Table 1a through 1e, the name of the author(s) and the number of the study are listed in column (1). Column (2) contains a short indication of the type of family therapy that is studied. Column (3) provides a more detailed description of the dependent measure that is used. In column (4), the population of patients is described, and in column (5) information is given about the sample size. The post-treatment effects and the follow-up effects can be found in columns (6) and (7); and, finally, the length of the follow-up period is listed in column (8). Tables 1f and 1g contain studies in which family therapy is compared to an alternative treatment, using behaviour ratings and recidivism respectively. In these two tables, column (1) and (2) again contain the

---

1. Weight $w = 2N / (8 + d^2)$; $d = (d_{post-treatment} + d_{follow-up}) / 2$
2. $Z_{combined} = \sum[Z_i / N]$; $Z_i$ = the Z-value computed for study $i$; $N$ is the number of studies; $Z_i = r_i \sqrt{N_i}$; $r_i = d_i / \sqrt{(4 + d^2)}$
3. Chi-square = $\sum[w(d_i - d_w)^2]$; $d_w$ = the weighted mean of the $d$-values; d.f. = $K - 1$, $K$ = the number of studies
4. Fail Safe N = $(\sum Z)^2 / 2.33 - N$; $N$ = the number of studies
name of the study and the type of family therapy studied, column (3) provides a description of the alternative treatment, in column (4) the dependent measure is described, column (5) provides information about the population of patients, and column (6) contains information about the sample size. The post-treatment and the follow-up effects can be found in columns (7) and (8), and the length of the follow-up period is listed in column (9).

The overall post-treatment effect could be computed over 10 studies. The average outcome of these ten studies is $d = 0.70$. This means that, immediately after treatment, subjects with a family therapeutic treatment score 0.70 standard deviations better on the dependent variables than those subjects who receive alternative, minimal, or no treatment. This implies that the average patient who received family therapy is better off than 76% of the patients in the various control groups. The follow-up effect could be computed over 13 studies, resulting in an average effect of $d = 0.55$. The mean follow-up period is one-and-a-half years.

The average effects sizes of post-treatment and follow-up were tested for statistical significance by Stouffer's Combined Test. The results of this test show that the post-treatment effect and the follow-up effect are both highly significant ($Z_c = 7.06$, $P < 0.001$ at post-treatment and $Z_c = 9.36$, $P < 0.001$ at follow-up).

Because no unpublished material could be included in the meta-analysis, a test was carried out to estimate the possibility of 'publication bias'. This test resulted in a 'Fail Safe N' of 82 studies at post-treatment, and 'Fail Safe N' of 197 studies at follow-up. This means that 82 studies with no difference between experimental and control groups would have to exist before our post-treatment effect would be rendered non-significant. At follow-up an even greater number of 197 studies with no experimental-control differences would have to be added to our set of 13 studies before the follow-up effect becomes non-significant. It seems unlikely that so many unpublished studies with no differences exist. Therefore it seems reasonable to conclude that our results are not simply an artifact of not having included unpublished manuscripts.

A validity threat to meta-analytic results is raised by the fact that the data of studies that differ substantially in design and population of patients are combined into one outcome. Care should be taken not to add 'apples and oranges' (Rosenthal, 1984). This can be tested by a chi-square test for homogeneity. If the set of outcomes of the various studies form a statistically homogeneous group, we can conclude that the studies were similar enough to be combined. If the various study outcomes form
**Table 1a** Studies comparing family therapy with a no-treatment control condition using behaviour ratings

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study number and authors</td>
<td>Type of family therapy</td>
<td>Dependent measure</td>
<td>Population of patients</td>
<td>N</td>
<td>Cohen d, post-treatment</td>
<td>Cohen d, follow-up</td>
<td>Number of months follow-up</td>
</tr>
<tr>
<td>1. Fisher (1978)</td>
<td>Non-behavioural</td>
<td>Parent evaluations of behaviour IP</td>
<td>Children with behaviour disorder</td>
<td>24.0</td>
<td>1.18</td>
<td>1.15</td>
<td>12.0</td>
</tr>
<tr>
<td>2. Santa Barbara (1979)</td>
<td>Behavioural</td>
<td>Teacher ratings behaviour IP; performance in school</td>
<td>Children with behaviour disorder</td>
<td>279</td>
<td>—</td>
<td>0.29</td>
<td>6.00</td>
</tr>
</tbody>
</table>

* In this study, pre- and postscores of the experimental group are compared ('within group' design). A control group was either not included in the original design or excluded from the analysis because it also contained some kind of family therapy.
### Table 1b: Studies comparing family therapy with a no-treatment control condition using recidivism

<table>
<thead>
<tr>
<th>Study number and authors</th>
<th>Type of family therapy</th>
<th>Dependent measure</th>
<th>Population of patients</th>
<th>N</th>
<th>Cohen $d$, post-treatment</th>
<th>Cohen $d$, follow-up</th>
<th>Number of months follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. Beal and Duckro (1977)</td>
<td>Behavioural</td>
<td>Number of closed court cases</td>
<td>Juvenile status offenders</td>
<td>98</td>
<td>0.52</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>5. Barton, et al. (1985)</td>
<td>Short term behavioural</td>
<td>Per cent that has to appear in court</td>
<td>Mild juvenile delinquents</td>
<td>27</td>
<td>—</td>
<td>1.15</td>
<td>13.0</td>
</tr>
<tr>
<td>6. Alexander and Parsons (1973)</td>
<td>Behavioural</td>
<td>Number of court cases</td>
<td>Mild juvenile delinquents</td>
<td>56</td>
<td>—</td>
<td>0.41</td>
<td>6–18</td>
</tr>
<tr>
<td>7. Klein et al. (1977)</td>
<td>Behavioural</td>
<td>Number of court cases of siblings</td>
<td>Mild juvenile delinquents</td>
<td>56</td>
<td>—</td>
<td>0.38</td>
<td>36.0</td>
</tr>
<tr>
<td>Study number and authors</td>
<td>Type of family therapy</td>
<td>Dependent measure</td>
<td>Population of patients</td>
<td>N</td>
<td>Cohen $d$, post-treatment</td>
<td>Cohen $d$, follow-up</td>
<td>Number of months follow-up</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------------------</td>
<td>-------------------</td>
<td>------------------------</td>
<td>----</td>
<td>--------------------------</td>
<td>----------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>8. Leff (1982)</td>
<td>Behavioural</td>
<td>Ratings of communication between family members</td>
<td>Adult schizophrenic patients</td>
<td>24</td>
<td>0.63</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>9. Parsons and Alexander (1973)</td>
<td>Behavioural</td>
<td>Experimenter ratings of audio recordings of adolescents with conduct disorder</td>
<td>Adolescents with conduct disorder</td>
<td>40</td>
<td>0.87</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>10. Slipp and Kressel (1978)</td>
<td>Behavioural</td>
<td>Patient satisfaction about therapy</td>
<td>Adults with individual and family complaints</td>
<td>19</td>
<td>0.48</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>11. Sigal (1976)</td>
<td>Long-term non-behavioural</td>
<td>Parents satisfaction about family functioning</td>
<td>Children with behaviour problems</td>
<td>93</td>
<td>—</td>
<td>0.04</td>
<td>55.0</td>
</tr>
<tr>
<td>12. Martin (1977)</td>
<td>Short-term behavioural</td>
<td>Experimenter ratings of other reports of problem behaviour IP</td>
<td>Children with behaviour problems</td>
<td>43</td>
<td>0.74</td>
<td>0.74</td>
<td>6.00</td>
</tr>
</tbody>
</table>
Table 1d: Studies comparing family therapy with a minimal-treatment condition using behaviour ratings

<table>
<thead>
<tr>
<th>(1) Study number and authors</th>
<th>(2) Type of family therapy</th>
<th>(3) Dependent measure</th>
<th>(4) Population of patients</th>
<th>(5) Cohen $d$, post-treatment</th>
<th>(6) Cohen $d$, follow-up</th>
<th>(7) Number of months follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>13. Garrigan and Bambrick (1977)</td>
<td>Short-term non-behavioural</td>
<td>Independent ratings of behaviour in school</td>
<td>Adolescents with emotive disorders</td>
<td>26</td>
<td>0.28</td>
<td>—</td>
</tr>
<tr>
<td>11. Sigal (1975)</td>
<td>Long term non-behavioural</td>
<td>Parents ratings of frequency and severity of the problems</td>
<td>Children with school, behaviour and other problems</td>
<td>93</td>
<td>—</td>
<td>0.12</td>
</tr>
</tbody>
</table>
Table 1c Studies comparing family therapy with a minimal treatment condition using recidivism

<table>
<thead>
<tr>
<th>Study number and authors</th>
<th>Type of family therapy</th>
<th>Dependent measure</th>
<th>Population of patients</th>
<th>N</th>
<th>Cohen $d$, post-treatment</th>
<th>Cohen $d$, follow-up</th>
<th>Number of months follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>8. Leff (1982)</td>
<td>Eclectic</td>
<td>Relapse of schizophrenic symptoms</td>
<td>Adult schizophrenic patients</td>
<td>24</td>
<td>0.92</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>(1) Study number and authors</td>
<td>(2) Type of family therapy</td>
<td>(3) Alternative treatment</td>
<td>(4) Dependent measure</td>
<td>(5) Population of patients</td>
<td>(6) Cohen d, post-treatment</td>
<td>(7) Cohen d, follow-up</td>
<td>(8) Number of months follow-up</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>--------------------------</td>
<td>---------------------</td>
<td>---------------------------</td>
<td>---------------------------</td>
<td>---------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>14. Glick et al. (1985)</td>
<td>Inpatient; behavioural</td>
<td>Inpatient individual therapy</td>
<td>General functioning of IP</td>
<td>Adults with cognitive and affective disorders</td>
<td>95</td>
<td>0.51</td>
<td>0.33</td>
</tr>
<tr>
<td>15. Langsley et al. (1968, 1971)</td>
<td>Behavioural</td>
<td>Hospitalization</td>
<td>Autonomous solving of daily problems</td>
<td>Adults referred to a clinic</td>
<td>300</td>
<td>0.28</td>
<td>6.00</td>
</tr>
<tr>
<td>16. Ro-Trock et al. (1977, 1980)</td>
<td>Inpatient, short-term; non-behavioural</td>
<td>Individual psychotherapy</td>
<td>Number of days needed after discharge for full recovery of functioning</td>
<td>Adolescent psychiatric patients</td>
<td>28</td>
<td>—</td>
<td>0.88</td>
</tr>
<tr>
<td>Study number and authors</td>
<td>Type of family therapy</td>
<td>Alternative treatment</td>
<td>Dependent measure</td>
<td>Population of patients</td>
<td>Cohen $d$, post-treatment</td>
<td>Cohen $d$, follow-up</td>
<td>Number of months follow-up</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------------------------</td>
<td>-----------------------</td>
<td>------------------</td>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>16. Ro-Trock et al. (1977, 1980)</td>
<td>Inpatient, short-term; non-behavioural</td>
<td>Individual psychotherapy</td>
<td>Rehospitalization; number of outpatient consults</td>
<td>Adolescent psychiatric patients</td>
<td>24</td>
<td>-</td>
<td>-1.09</td>
</tr>
<tr>
<td>17. Falloon et al. (1985)</td>
<td>Behavioural</td>
<td>Individual psychotherapy</td>
<td>Clinicians ratings of schizophrenic symptoms</td>
<td>Adult schizophrenic patients</td>
<td>36</td>
<td>1.39</td>
<td>1.46</td>
</tr>
<tr>
<td>18. Barton et al. (1985)</td>
<td>Behavioural</td>
<td>?</td>
<td>Number of placements in foster homes</td>
<td>Children with behaviour disorder</td>
<td>274</td>
<td>-</td>
<td>1.15</td>
</tr>
<tr>
<td>19. Barton et al. (1985)</td>
<td>Behavioural</td>
<td>Placement in half-way homes</td>
<td>Number of criminal charges</td>
<td>Severe juvenile delinquents</td>
<td>74</td>
<td>-</td>
<td>0.87</td>
</tr>
</tbody>
</table>
a heterogeneous group, we have to conclude that the combined studies probably were too different to be combined. The 10 outcomes that were combined into one post-treatment effect size are homogeneous (chi-square = 8.45, d.f. = 9, P < 0.55), indicating that the various studies were similar enough to be combined. However, the 13 combined study outcomes for the follow-up effect are heterogeneous (chi-square = 55.81, d.f. = 12, P < 0.001). This suggests that these studies are too diverse in methodology or population of subjects to allow a combination of their results. Therefore, our follow-up result cannot be considered to be a valid estimate of the real treatment effect until one or more factors causing this heterogeneity are specified.

**Exploration**

Finding the variables responsible for the heterogeneity of a set of results can reveal important moderating and mediating effects (Baron and Kenny, 1986). We first tried to explain the heterogeneity of the 13 follow-up outcomes by the fact that these studies varied greatly in the population of patients that were treated [see Tables 1a to e, column (4), and Table 1f and 1g, column (5)]. However dividing the follow-up studies into three groups of symptoms (delinquency, conduct disorder and schizophrenia) does not lead to homogeneity of results within these groups. Categorizing the follow-up studies on the basis of the length of their follow-up period, does however, lead to homogeneity of results within each category of studies. Not too surprisingly then, the length of a follow-up period appears to influence the treatment effect.

![Figure 1. Family therapy effectiveness at various lengths of follow-up.](image-url)
Calculation of the ‘correlation ratio’, a method that takes the possibility of non-linearity of a relation into account (Hays, 1974) results in a correlation of 0.66 between the two variables ‘length of follow-up’ and ‘effect size’. Figure 1 displays the relationship to which the correlation refers.

At a follow-up of 6 months, a ‘moderate effect’ of $d = 0.41$ is measured ($N = 4$), that increases into a ‘big effect’ at a follow-up of 10 months ($d = 1.14$, $N = 3$). Eighteen months after therapy, the effect seems to decline, but still remains ‘big’ ($d = 0.84$, $N = 4$). After 18 months, then, the effect steadily drops. ($d = 0.40$, $N = 1$, follow-up of 2 years) and seems to disappear after 4.5 years ($d = 0.08$, $N = 1$, follow-up of 55 months)—though we have only two studies that extend their evaluation past 18 months. We see that the correlation ratio refers to a strong curvilinear relation between ‘length of follow-up’ and ‘effect size’, where the effect increases in the first year after therapy and drops down again after 18 months. This curvilinear relation explains 40% of the variance in follow-up outcomes. Because of the small number of studies, however, these data can only tentatively suggest the possibility of such a curvilinear relation between the two variables.

Discussion

According to the results of our meta-analysis, immediately after treatment family therapy has a ‘moderate effect’ compared to other treatments, minimal treatment or no treatment (Hazelrigg et al., 1987) also found a ‘moderate effect’ in their meta-analysis, although the exact outcome was somewhat lower than in our meta-analysis ($d = 0.46$). This variance in outcomes between the two meta-analyses can partially be attributed to the fact that the two meta-analyses included a different set of studies. There was an overlap of only 53% in included studies between the two meta-analyses. The difference in outcome of the two meta-analyses can also be influenced by differences in procedures and selection strategies of data when several results in a study were available (Rosenthal, 1984). For follow-up, Hazelrigg et al. report a correlation of $r = -0.30$ between ‘length of follow-up’ and ‘effect size’. In our meta-analysis, a curvilinear relation is tentatively suggested between these two variables. Thus, our findings yield an interesting new suggestion about a possible curvilinear course of the follow-up effect of family therapy.

In a meta-analysis of 400 outcome studies, Smith and Glass (1977) estimated the effectiveness of psychotherapy in general. They found an
average effect of $d=0.68$ after 3.75 months. For follow-up, Smith and Glass did not report a separate effect size. However, they did assess, a correlation of $r=-0.10$ between 'length of follow-up' and 'effect size'. If we compare our post-treatment results of $d=0.70$ (0–2 months after therapy) with the results of Smith and Glass (3.75 months after therapy), we can conclude that family therapy is as effective as other psychotherapies shortly after treatment but not more effective.

For a comparison with alternative treatments at follow-up, we may inspect the results of the six follow-up studies in Tables 1f and g, where family therapy was compared with other therapies. (study numbers 14 to 19 in Tables 1f and g). Family therapy seems to have a bigger effect 10 months after treatment ($d=0.67$, s.d. $=0.64$, $P<0.001$; Fail Safe $N=61$), yet this effect is highly variable across the various studies (chi-square $=37.72$, d.f. $=5$, $P<0.01$). Although the variance of these follow-up results is substantial, the effect of family therapy is bigger than that of other therapies in five of the six studies. This suggests that family therapy may be more effective than other treatments one year after the end of treatment. More comparative data are needed before the size and conditions of this effect can be adequately assessed.

Our follow-up results suggest that the effect of family therapy increases during the first year after therapy, and that it stays large until 18 months after therapy. After more than 18 months the effect may drastically decrease. However, we only have two studies with which to show this decline. More studies with extended follow-up periods are needed before the long-term effectiveness of family therapy can be evaluated.

**Acknowledgements**

The authors are grateful to Dr Harrie Vorst of the Section Methodology at the Department of Psychology of the University of Amsterdam, for his help with the analyses of the data.

This research was done as a senior thesis by the first author under supervision of the other two authors.

**References**


