

Psychotherapy for sexually abused girls: psychopathological outcome findings and patterns of change

J. TROWELL, I. KOLVIN, T. WEERAMANTHRI, H. SADOWSKI,
M. BERELWITZ, D. GLASSER and I. LEITCH

Background Controversy exists about the efficacy of psychotherapy for the mental health problems of sexually abused children.

Aims To compare the relative efficacy of focused individual or group therapy in symptomatic sexually abused girls, and to monitor psychiatric symptoms for persistence or change.

Method A multi-centre psychotherapy outcome study recruited 71 sexually abused girls aged 6–14 years who were randomly assigned to focused individual psychotherapy (up to 30 sessions) or psychoeducational group therapy (up to 18 sessions). Changes over the course of the study were monitored.

Results Both treatment groups showed a substantial reduction in psychopathological symptoms and an improvement in functioning, but with no evident difference between individual and group therapy. However, individual therapy led to a greater improvement in manifestations of post-traumatic stress disorder (PTSD).

Conclusions The beneficial effects on PTSD support the use of individual therapy. However, the small sample size and lack of a control group limit conclusions about changes attributable to treatment.

Declaration of interest None. Funding described in Acknowledgements.

Child abuse and child sexual abuse (CSA) are important risk factors for subsequent mental health problems. Although considerable efforts have been devoted to the identification and investigation of abuse, concern has been expressed about the limited treatment resources for dealing with its mental health consequences and the lack of knowledge as to which treatments are most effective, for whom they should be provided, and the symptoms or problems that are most likely to be influenced by treatment.

Follow-up studies of sexually abused children have investigated both short-term and long-term effects. Beitchman *et al* (1991) reviewed 42 studies for short-term effects with the following conclusions:

- (a) sexually abused children are more likely to develop some form of inappropriate sexual and disturbed behaviour than those not abused;
- (b) the greater the frequency and duration of sexual abuse, the more pronounced the effects;
- (c) CSA involving force and/or penetration, or sexual abuse perpetrated by the child's biological father or stepfather, is associated with poorer outcome and greater psychological problems;
- (d) sexually abused children are more likely than those not abused to come from families with a higher incidence of marital separation or divorce, parental substance misuse and psychiatric disorder.

Cotgrove & Kolvin (1996), drawing on literature reviews (Zanarini *et al*, 1989; Brown & Anderson, 1991; McLeer *et al*, 1992) and research (McLeer *et al*, 1994), concluded that there are five main long-term associations with CSA. These are *psychological symptoms*, consisting of depression, anxiety, low self-esteem, guilt, sleep disturbance and dissociative phenomena; *psychiatric disorders*, particularly

depression and anxiety, eating disorders and *borderline personality disorder* in adulthood; *problem behaviours*, including self-harm, drug use, sexual behaviour problems, running away and conduct disorder; and *social relationship problems*, such as social withdrawal, sexual promiscuity and re-victimisation. It was recognised that these problems were more likely to be reported by clients in long-term therapy or those referred to psychiatric or social services. This research has demonstrated widespread and serious disorders compounded by much comorbidity in those referred for psychotherapy (Trowell *et al*, 1999; further details available from the author upon request).

Stevenson's wide-ranging review of psychotherapy outcome (Stevenson, 1999) is not optimistic regarding the efficacy of treatment of the mental health problems of sexually abused children. Nevertheless, it has been argued that, given the diverse and enduring effects of sexual abuse on the mental health of some sexually abused children, psychological treatment may benefit those with symptoms by ameliorating or preventing medium- and long-term effects (Finkelhor & Berliner, 1995; Jones & Ramchandani, 1999).

METHOD

In this study, sexually abused children were randomly allocated to individual or group psychotherapy, with parallel support for their carers. Outcome measures were chosen to assess psychiatric disorder, global impairment of functioning, and three post-traumatic stress disorder (PTSD) dimensions, namely re-experiencing of traumatic events, avoidance of stimuli and increased arousal. The hypotheses to be tested were:

- (a) Psychoeducational group therapy, in comparison with individual psychotherapy, would lead to a greater reduction of (or recovery from) psychiatric disorders, reduce the frequency of sexualised and eroticised behaviour (Yates, 1982), and promote normal emotional development.
- (b) Those receiving psychoeducational group therapy would benefit more in the short term than would patients receiving individual psychotherapy, whereas the latter would benefit more from the treatment in the long term.
- (c) The two treatments would have differential effects on child psychiatric

disorders and PTSD, with regard to both the reduction of (or recovery from) psychiatric disorders and the traumatogenic responses.

Study design

A 'comparison' design with random allocation rather than a control design was used, as it was considered unethical to leave untreated a control group of traumatised, sexually abused girls for the duration of the study. At the end of the control research assessment, the girls and their carers were informed which specific therapies were available.

The treatments were carried out by trainee psychotherapists or experienced mental health professionals with the guidance of a manual and with skilled clinical supervision. The two main institutions involved were a tertiary centre (Tavistock Clinic) in north London and a community clinic (Camberwell Child and Adolescent Service) in south London. It was decided to offer either individual psychotherapy in the form of brief, focused psychoanalytic sessions once weekly for up to 30 sessions, or group psychotherapy (comprising both psychotherapeutic and psychoeducational components) for up to 18 sessions. Both therapies were complemented by supportive work with the carers, helping to ensure the child's attendance for therapy and enabling the carers to address any emergent family issues. The age group 6–14 years was selected because by the age of 6 years most girls would be able to read and to cope with the assessment instruments, and the oldest children would still be in school and would thus be available for a second year of follow-up. Variation in delivery across therapists and centres was minimised by use of manuals and by close supervision by experienced therapists. Theoretically, this should be complemented by process analysis of recorded sessions (Hardy & Shapiro, 1985) but such methods were considered not to be appropriate for sexually abused children.

Improvement has been shown to be related to the duration of therapy, being proportionately greater in early sessions but progressing more slowly as the number of sessions increases (Howard *et al.*, 1986). This has led to the suggestion that in comparisons of two forms of therapy that are intrinsically different, the duration of therapy should be controlled. In this study, although there were differences in the

numbers of sessions for the two therapies, the length of each session was such that the face-to-face contact time was approximately the same. Frequency and duration issues are discussed elsewhere (Trowell *et al.*, 1995).

Inclusion and exclusion criteria

The entry criteria for this study were as follows:

- (a) contact sexual abuse had occurred on the basis of balance of probabilities verified by social services and/or court procedure;
- (b) the children involved were school-aged girls (6–14 years of age);
- (c) consent to participate in the study had been given by the child and the child's legal guardian;
- (d) symptoms of emotional or behavioural disturbance warranting treatment were present;
- (e) the abuse had been disclosed within 2 years prior to referral, regardless of when the abuse had actually occurred.

Symptoms of emotional or behavioural disturbance were not tightly defined. It was anticipated that there would be a description, by professional staff who had referred the children, in similar terms to those used in referrals to mental health clinics. This proved to be the case, except that the symptoms proved to be wider and more severe than in children referred to community clinics (Trowell *et al.*, 1999). Systematisation of the psychopathological diagnosis and coding of impairment were related to the subsequent baseline assessment.

Exclusion criteria were severe developmental delay; psychosis; lack of reasonable confidence that further abuse would not occur; and necessity for hospitalisation at the time of initial evaluation. On a case-by-case basis, other clinical and legal issues were taken into consideration.

Assessment

The children were assessed at baseline prior to therapy (a detailed account of the baseline assessments is available from the author upon request), a year after the start of therapy and at second-year follow-up. The carers were interviewed to gather as much information as possible about the child, to learn about the carers' responses to the abuse and to obtain information

about the carers' significant social history and abuse background. This paper focuses on psychopathology based on two assessment measures: the Schedule for Affective Disorders and Schizophrenia for School-age Children, known as Kiddie-SADS (K-SADS; Chambers *et al.*, 1985; Ambrosini *et al.*, 1989) and Orvaschel's scales for assessing PTSD (Orvaschel, 1989). In addition, the K-SADS (1986 revised version) provides data for coding a global assessment of impairment of functioning: the Kiddie Global Assessment Scale (K-GAS; Chambers *et al.*, 1985).

K-SADS

A shortened version of the standardised, semi-structured K-SADS diagnostic interview was used. This is a valid and reliable schedule that has been widely employed with children aged 6–17 years, and has proved to be a reliable measure of depression, anxiety, emotional disorder and conduct disorder (Kaufman *et al.*, 1997) (see Appendix).

Global impairment of functioning

Assessment of global impairment of functioning (concerning social, psychological or school functioning) was achieved using the K-GAS (1986 version). This is based on the Children's Global Assessment Scale (C-GAS) (Shaffer *et al.*, 1983). Impairment is rated on a scale of 1 to 9 (1–3, major impairment; 7–9, little impairment). In recent research using kappa (Cohen, 1960; G. Kolaitis & T. Korpa, personal communication, 2001) satisfactory reliability has been reported (see Appendix).

Post-traumatic stress disorder

The 1989 version of Orvaschel's PTSD scale was used, which is an extension of the K-SADS instrument (Orvaschel, 1989). The content is similar to the criteria outlined in DSM-IV (American Psychiatric Association, 1994), comprising 19 items with four domains: re-experiencing of trauma, inappropriate sexual behaviour, persistent avoidance of stimuli, and persistent increased arousal. The data are obtained by interviewing the child, apart from information on inappropriate sexual behaviour which (according to Orvaschel) is more reliably obtained from the child's carer. (Normally, inappropriate sexual behaviour is not collected as part of a PTSD profile; however, it has been included as

such by Orvaschel.) The PTSD symptoms were coded in a number of different ways: for most analyses, a bimodal system of coding of 'present' or 'absent' was employed; when studying a change, in order to provide a wider range of scores, the three-point coding system was used; sub-scale or dimension scores were obtained by summation of the relevant items; and PTSD as a category was coded as present (1) or absent (0).

Analyses

The study was planned to compare two active treatments. There is some scientific evidence in the literature for the efficacy of group therapy in sexually abused children, but little supporting psychodynamic individual therapy. In a study of therapy with psychiatrically disturbed children using an appropriate criterion of outcome (Kolvin *et al.*, 1981, 1988), good clinical outcome was found in about 75% of children who had been given group therapy and in 47% of those whose parents were given counselling; this difference proved statistically significant ($P < 0.01$). On this basis, in order to detect such a difference between the two treatment programmes at follow-up with a significance level of 5% (two-tailed test) and 80% power, 44 patients should have been recruited for each treatment group; in addition, a small allowance needed to be made for drop-outs. Initially, therefore, we planned to have 48 participants in each treatment programme, with six children in each group. In practice, the group sizes proved smaller: on this basis, the total of 96 participants was considered to be an overestimate, and 90 entering therapy would suffice.

Outcome

Descriptive statistics were used to provide a baseline demographic picture of the two treatment groups at the point of entry to the study. Categorical data were analysed by chi-squared tests or comparison by odds ratios with 95% confidence intervals (CIs). Parametric data of outcome were explored using effect sizes. An account of effect size and its usefulness in child psychotherapy research has been provided by Kazdin (1998, 2000). The term refers to the magnitude of difference between two conditions or groups and is expressed in standard deviation units (Kazdin, 1998, 2000), with standard deviation being that of the pooled mean score data at baseline. Kazdin also

states that in research an effect size is a measure of the magnitude of the effect of a relation between two conditions and is distinguished from statistical difference – and these concepts are quite different. A useful but arbitrary guide to interpreting effect size is provided by Cohen (1977, 1992) as follows: small 0.20, medium 0.50, and large 0.80.

Meta-analysis of child psychotherapy data has demonstrated that the average outcome of child psychotherapy is slightly more than two-thirds (effect size about 0.70) of a standard deviation better than that of untreated control children (Smith & Glass, 1977; Casey & Berman, 1985; Kazdin, 1998, 2000). However, alternative models of active therapy tend to be equally effective (Kazdin, 1998, 2000). As we had no untreated control group, we followed Cohen's guide of an effect size of 0.50 for a medium effect. In this we were supported by the findings of Fine *et al.* (1991) when comparing active child therapies. However, effect size does not make any allowance for the influence of any relevant covariates, and analyses of covariance were therefore undertaken. With the small sample size only a few covariates were chosen: a relevant baseline index of impairment; an index of severity of PTSD dysfunction at baseline; an index of compliance as represented by therapy sessions; and an index of severity of abuse at baseline. Three covariates were considered to be the maximum number that could be used. A common covariate measure for all covariance analyses was the index of impairment. However, where the outcome was not this measure, then the baseline score on the PTSD-dependent measure was used in a further covariance analysis as an index of severity at baseline.

Changes

Patterns of change across time are demonstrated by cross-tabulations of relevant categories. Magnitude of change irrespective of the treatment group was explored using analysis of variance. In addition, the magnitude of reduction of the mean scores from baseline to the two follow-up points is indicated in terms of standard deviations.

CONSORT model

Subjects

A CONSORT model (Begg *et al.*, 1996) is shown in Fig. 1. A total of 94 participants

were recruited initially from the authors' own clinics and from professional agencies in the community in north and south London, but the process proved slower than anticipated; 13 children proved ineligible, on the basis of information in the referral letters or telephone contact, or where there was unwillingness to participate in initial assessments. Owing to the absence of interview data for these 13 children, it was not possible to compare them with the trial participants. In all, 81 girls were interviewed, and of these 10 proved unsuitable for diverse reasons – because they were not in a safe family situation, or could not or would not commit themselves to regular attendance, or had been sexually abused but had little evidence of symptoms. One child objected to random allocation, as she felt unable to cope with the perceived intimacy of individual therapy. Four children were allocated randomly but did not attend. A total of 71 children entered therapy. These numbers are linked to available K-SADS and global assessment of impairment data. The number of participants with completed PTSD interviews was less than 71; this was because a few girls at the baseline refused to talk about the abuse and its aftermath.

Allocation to treatment regimen and random allocation

The north London centre recruited 51 participants and the south London centre recruited 20. North London was responsible for random allocation to group and individual therapies. In north London, with its relatively high recruitment, the system worked well, although there were a few inevitable delays while awaiting a group to be constituted; however, in south London the delays proved greater. The group therapy supervisors decided that, in general, to have twins or sisters in the same group would be unhelpful to the group. Because of the number of abused siblings and twins, there were practical problems that had to be solved on a case-by-case basis to ensure attendance.

Consent

Prior to the baseline assessment, the nature of the project was explained to each child and her carer. This included an account of the duration of the treatment offered, the nature of the assessment before and after treatment and at second-year follow-up, and the randomisation. The children's

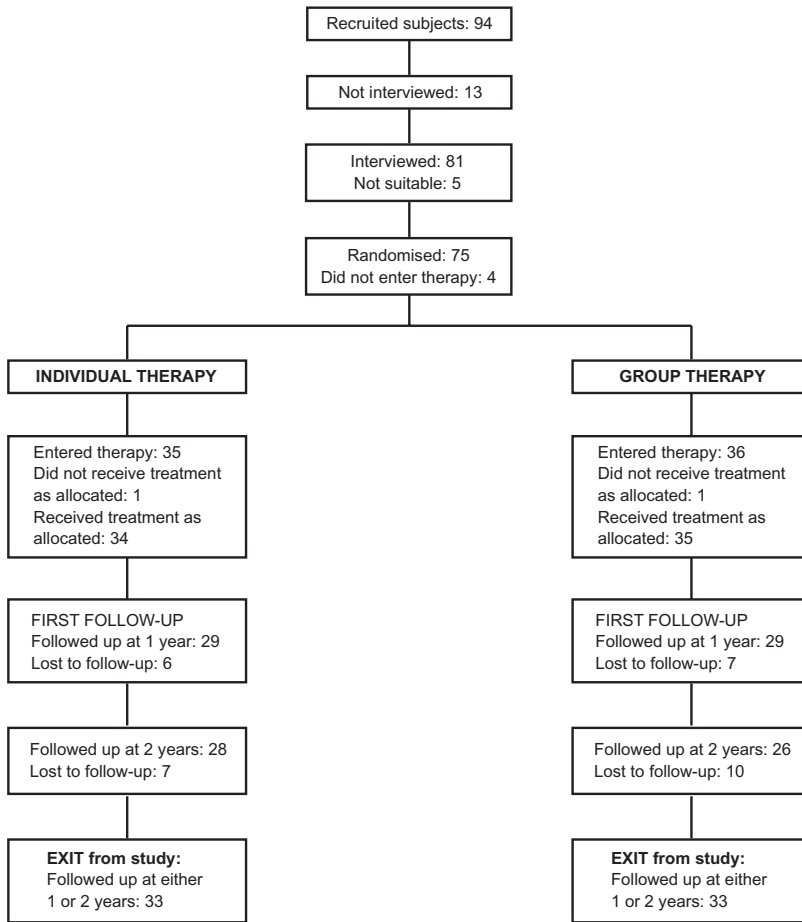


Fig. 1 Modified CONSORT diagram: psychiatric disorders and global assessment of impairment data.

carers were invited to commit themselves to the trial, and if the child was old enough she was invited to assent to the treatment package. If the legal carer or child did not wish to be part of the project, access was available to the ordinary resources at the referring clinics and hospitals; there were few refusals. At the end of the baseline assessment the allocation slot was revealed, and the families were invited to sign the consent forms and the child the assent forms.

Dropouts, losses and compliance with therapy

The characteristics of the participants are shown in Table 1. In addition, the patients' progress within the trial is illustrated in a modified CONSORT diagram. Of the final 71 patients, 35 were allocated to individual therapy and 36 to group therapy. The mean percentage of attended sessions for patients receiving individual therapy was 88% (s.d.=24.9, 95% CI 79.3–96.7); in group therapy, the mean percentage of attended sessions was again 88% (s.d.=21.6, 95% CI 80.7–95.3). Of those who did not complete the full course of therapy, 'completers' were defined as those who had attended at least 30% of the defined sessions for individual or group therapy.

Of the patients entering individual therapy, one did not receive treatment as

Table 1 Demographic characteristics of the study sample

	Individual therapy		Group therapy		χ^2	P
	n	%	n	%		
The child						
Under 10 years old	17/34	49	17/36	47	0.00 ¹	NS
Attends secondary school	8/32	25	11/36	31	0.06	NS
Above average academically	23/33	70	24/36	67	0.00	NS
Has a known best friend	21/31	68	23/33	70	0.00	NS
Pre-school separation for longer than 3 months	7/31	23	5/31	15	0.19	NS
Legal order	13/33	39	14/34	41	0.00	NS
Abused by her parent	12/35	34	18/36	50	1.21	NS
More than one abuser	15/35	43	13/36	36	0.11	NS
More than 10 abuse incidents	22/35	63	17/36	47	1.18	NS
More than 2 years' duration	14/35	40	13/36	36	0.01	NS
Sexual abuse						
Touching with or without clothes	8/35	23	4/36	11	6.0 ²	NS
Genital touching or simulated intercourse	12/35	34	15/36	42		
Oral, anal or vaginal penetration	15/35	43	17/36	47		

NS, not significant.

1. Child data values 1 d.f.

2. Sexual abuse values 4 d.f.

defined (she attended less than 20% of sessions) and was considered to be a drop-out. However, most (29) received substantial treatment (i.e. attended 80% or more of the defined sessions), and five received adequate treatment, that is 30% or more of the defined sessions. Of those entering group therapy, one did not receive treatment as defined, most (29) received substantial treatment and six received adequate treatment. Thus, prior to the first follow-up, 34 children in individual therapy and 35 in group therapy received at least adequate treatment.

However, it will be noted that there were only 58 participants with K-SADS data at the first follow-up; thus, there were more completers than those who were assessed at follow-up. There were a number of reasons for this. Some patients declined assessment as they did not wish to be reminded of their traumatic experience; some families had moved and it proved difficult to track them down. Although much time was spent on encouraging families to attend, the resources were not always available to arrange repeated home visits or multiple clinic appointments. Social service departments felt that the children should not be pursued for follow-up. The net result was that although 34 families had received adequate treatment in the individual therapy sessions and 35 in the group therapy sessions, only 29 families in each of the treatments were assessed at first-year follow-up, and 66 at exit from the study. Losses on the PTSD measure were greater than on the K-SADS; for instance, 68 girls with PTSD data entered therapy (34 individual and 34 group); 56 attended the first-year follow-up (28 individual and 28 group); 49 attended the second-year follow-up (27 individual and 22 group); and 64 were available at exit (32 individual and 32 group). Some children refused to answer PTSD questions, they would not discuss the trauma.

It was not necessarily the children who missed the first follow-up who missed the second: some of those who missed the first follow-up had been encouraged to attend the second, and a number of children who attended the first follow-up did not attend the second. This compounded the problems for multivariate analyses, for which a complete set of data is necessary. To deal with missing values, we used the technique of replacing missing values at the second follow-up with the last available measure, which is the notion of the last observation carried

forward (Diggle, 1998): the *exit notion*. As shown by the modified CONSORT diagram (Fig. 1), this technique gave rise to a larger sample ($n=66$ with K-SADS data) at the second-year follow-up (previously $n=54$).

The therapy

Individual therapy plus carer support

Children allocated to individual therapy were seen once weekly for up to 30 sessions. The sessions lasted 50 minutes and took place at the same time on the same day each week, with the same therapist in the same room. Carers were seen by social workers (Rushton & Miles, 2000) in parallel, generally every 2 weeks; some carers chose not to attend, but most came for some sessions. The children were provided with play materials. The topics to be addressed were listed in the manual; the topics arose as they occurred to the girl, but the therapists and their supervisors ensured that over the time all the topics were raised. A checklist was completed after each session. The early sessions (usually the first five) were the engagement phase; the next 15 sessions were focused on the issues identified as relevant to that particular child; the final 10 sessions focused on separation and ending, while reworking any key topics that had emerged. Individual treatment was supervised every other week.

Group therapy plus carer work

The therapy groups, which were psycho-educational as well as psychotherapeutic, were led by co-therapists. All except one group had both a male and a female leader. There were up to 18 group sessions, involving generally five girls. Each group had a prearranged topic, the girls had notebooks and task sheets, and there were also play materials. Each session began with news and then the topic for the week; information and suggestions were given and explained in the group. The group also had time to tackle the relationships between the girls and their relationship with the co-therapists, and this was linked to past and current relationships, losses and disruptions. The groups were intense and the co-therapists required supervision after each session. Carers were worked with alongside the groups by carer workers, mostly separately. Carers' groups had some difficulties (Rushton & Miles, 2000), although some functioned well. However,

when the carer groups were a mixture of parents and substitute carers, the members did not have a great deal in common.

Outside agencies

In many cases, local authority social workers supported the attendance by visiting the homes, paying fares, or bringing the family. Liaison with the schools was also important, as was attendance at case conferences and writing review reports. This work was undertaken by the team, not by the child therapists (group or individual).

Intention to treat

The participants in this research were included not because they belonged to a specific diagnostic category but because they presented with acute and chronic psychological symptoms following contact sexual abuse; they frequently suffered from multiple disorders. The therapy was intended to treat symptoms rather than specific disorders or diagnoses – both treatments had generic and abuse-specific components as advocated by Jones & Ramchandani (1999). As identified by Guthrie (2000), the focus of the work included engaging ambivalent girls and their carers. Where the treatment touched painful or difficult areas the work included maintenance of the therapeutic alliance, prevention of drop-outs by work with the carers, referrers, social workers and schools as appropriate, and management of anxieties and appropriate handling of the post-abuse and current concerns as they arose. This necessitated other clinical professionals working with the carers and outside agencies so that the therapists could focus on the generic issues and the multi-symptom complaints. Coordinating all the work and ensuring that the time available for therapeutic work was protected required considerable staff resources (see Trowell & Kolvin, 1999).

RESULTS

Background features

The majority of the girls (45/71=63%) were White; 8 (11%) were Black Caribbean; 7 (10%) were of mixed parentage; 5 (7%) were Chinese; 4 (6%) were of Mediterranean origin, and 2 (3%) were of unknown origin. The assessed girls had a mean age of 10 years (s.d.=2.2, 95% CI 9.6–10.6). In all, 19 girls had been placed

Table 2 Kiddie Global Assessment Scale (impairment measure): change in scores by therapy

	n	Score		
		Mean	s.d.	95% CI
<i>Baseline assessment</i>				
All cases	71	5.01	1.20	4.73–5.30
Individual therapy	35	5.14	1.37	4.67–5.62
Group therapy	36	4.89	1.01	4.55–5.23
<i>First-year follow-up</i>				
Scores	58	6.40	1.36	6.02–6.74
Change from baseline				
All cases	58	1.43	1.46	1.05–1.82
Individual therapy	29	1.48	1.57	0.88–2.08
Group therapy	29	1.38	1.37	0.86–1.90
<i>Second-year follow-up</i>				
Scores	54	6.72	1.17	6.40–7.04
Change from baseline				
All cases	54	1.57	1.41	1.19–1.95
Individual therapy	28	1.54	1.37	1.00–2.07
Group therapy	26	1.62	1.47	1.02–2.21
<i>Exit from study</i>				
Scores	66	6.70	1.16	6.41–7.00
Change from baseline				
All cases	66	1.70	1.49	1.33–2.06
Individual therapy	32	1.68	1.55	1.13–2.25
Group therapy	34	1.71	1.45	1.20–2.21

High scores represent least impairment, low scores represent most impairment. Effect sizes are not recorded because they were low at both the first-year and second-year follow-ups.

in foster families, and about a third of these girls had lived with their new families for longer than a year; three had been placed in children's homes; and 49 girls were living in their families of origin. Those in the last group were categorised as 'children remaining in family', whereas the 22 children who were placed in either foster families or group homes formed the 'children looked after' group. The mean age of the girls allocated to individual therapy was 9.7 years (s.d.=2.4, 95% CI 8.9–10.5) and for those in group therapy it was 10.4 years (s.d.=2.0, 95% CI 9.8–11.1). On average, the girls removed from their families were about a year younger (9.5 years) than the girls remaining at home (10.7 years), and this difference was statistically significant ($t=2.16$, d.f.=1, $P<0.05$). For the purpose of analysis in this study, girls who had experienced CSA were allocated to two age groups: those under the age of 10 years formed the younger age group ($n=34$), whereas girls at least 10 years old were allocated to the older age group ($n=37$). Only 31 (63%) of the girls who remained with their families after the sexual abuse

had a father figure living at home at the time of the baseline assessment, and this was also true for nine (41%) of the girls who were placed in foster families or group homes after the abuse. In addition, girls who were 'looked after' were significantly more likely to live in placements with three or more children (18/22) than the girls who had stayed in their family of origin (29/49) at the time of the assessment ($\chi^2=5.7$, d.f.=1, $P<0.05$). Most of the sexually abused girls (39/49; 80%) who had remained in their family of origin – but only one fostered girl – were not subject to any legal order. All the girls had been investigated for sexual abuse by social services (Table 1); however, some abusers had either left the family or had been imprisoned, and for these no order was deemed necessary to ensure the girls' safety. The background features at baseline for the two treatment groups are shown in Table 1. Full information was not always available in these complex families. There is no significant difference between these groups. In contrast, but according to expectation, comparison of the 'looked after' and

'at home' groups revealed significant differences. Fewer (1/22=5%) of the looked after girls attended secondary school than did those at home (18/49=37%) ($\chi^2=13.8$, $P<0.001$); more had had more than one abuser (18/22=82% *v.* 10/49=20%; $\chi^2=21.5$, $P<0.001$); and more of the 'looked after' group had experienced abuse of more than 2 years' duration (13/22=59% *v.* 14/49=29%; $\chi^2=4.8$, $P<0.05$).

Outcome

At baseline assessment the main DSM-IV diagnoses for the 81 children evaluated were as follows: PTSD (73%); major depressive disorder (57%); general anxiety disorder (37%); separation anxiety disorder (58%) (Trowell *et al.*, 1999; further details available from the author upon request).

Univariate analyses

Comparison of therapy groups. Four measures (the K-GAS impairment measure and the three PTSD dimensions) were subjected to univariate analyses (sexual behaviour will be reported elsewhere). Means, standard deviations and change scores from baseline with effect sizes were also studied. An effect size of 0.5 (following Cohen, 1992) was used as the criterion of medium effect. The findings are noted in Tables 2 and 3. On the K-GAS impairment measure (Table 2) the between-group effect sizes never achieved 0.5. The same proved true for the PTSD dimension of 'persistent symptoms of increased arousal' (not shown in the tables). These latter data were not subjected to further analysis.

Table 3 indicates that, at baseline, there were no significant differences between the mean scores on the two PTSD dimensions for the two groups of active therapy. However, at follow-up, the mean change scores on individual therapy in comparison with group therapy were associated with effect sizes of 0.5 or greater, which represents at least a medium effect in favour of individual therapy, with the exception of baseline to second-year follow-up and to exit in the case of the 'persistent avoidance' dimension; therefore, five of the six effect sizes listed in Table 3 meet Cohen's criterion of medium effect and one suggests a large effect.

Comparison of 'at home' and 'looked after' groups. There is one difference at the first-year follow-up with effect size on the

Table 3 Severity of post-traumatic stress disorder and persistent avoidance of stimuli: mean change scores by therapy

Symptom	n	Score			Effect size ¹
		Mean	s.d.	95% CI	
<i>Re-experience of traumatic event</i>					
Baseline scores					
All cases	68	7.4	2.19	6.9–8.0	
Individual therapy	34	7.9	2.14	7.1–8.6	
Group therapy	34	7.0	2.18	6.2–7.7	
Increase from baseline at first-year follow-up					
All cases	56	1.16	2.5	0.48–1.83	
Individual therapy	28	1.82	2.4	0.89–2.75	0.60
Group therapy	28	0.50	2.5	0.49–1.48	
Increase from baseline at second-year follow-up					
All cases	49	1.18	2.32	0.52–1.85	
Individual therapy	27	1.96	2.08	1.14–2.79	0.79
Group therapy	22	0.22	2.28	0.79–1.24	
Increase from baseline at exit from study					
All cases	64	1.06	2.54	0.43–1.70	
Individual therapy	32	1.78	2.43	0.90–2.66	0.65
Group therapy	32	0.34	2.48	0.55–1.24	
<i>Persistent avoidance of stimuli</i>					
Baseline scores					
All cases	68	10.1	1.67	9.7–10.5	
Individual therapy	34	10.2	1.56	9.7–10.8	
Group therapy	34	10.0	1.80	9.4–10.6	
Increase from baseline at first-year follow-up					
All cases	56	2.1	2.5	1.4–2.7	
Individual therapy	28	2.6	2.2	1.8–3.5	0.66
Group therapy	28	1.5	2.7	0.44–2.6	
Increase from baseline at second-year follow-up					
All cases	49	1.8	2.1	1.2–2.4	
Individual therapy	27	2.1	2.3	1.2–3.0	0.36
Group therapy	22	1.5	1.8	0.72–2.3	
Increase from baseline at exit from study					
All cases	64	2.0	2.5	1.4–2.6	
Individual therapy	32	2.5	2.4	1.6–3.3	0.60
Group therapy	32	1.5	2.6	0.6–2.5	

High scores represent most impairment.

1. Individual therapy v. group therapy.

K-GAS score. The mean score of the 'at home' group at the first-year follow-up was 6.59 (s.d.=1.31, 95% CI 6.16–7.02); for the 'looked after' group the mean score was 5.95 (s.d.=1.39, 95% CI 5.28–6.62) (higher scores imply less impairment) with an effect size of 0.53. No other effect size achieved the criterion level.

Multivariate analyses

Analysis of covariance: change scores. Two-way analyses of covariance were undertaken, using as separate main effects (a) the

treatment category (individual or group) and (b) the placement of the child, either at home or away from home; and adjusting for three covariates. The first of these was the K-GAS measure of global impairment at baseline (log transformed) or, where focusing on PTSD, the impairment score (log transformed); in addition, a further analysis was undertaken using the specific PTSD dimension at baseline (log transformed). The second was an index of the number of therapy sessions. The third was an index of severity of abuse.

The analyses were based on measures of change available at three time points: baseline assessment to first-year follow-up; baseline to second-year follow-up; and finally baseline to exit, which incorporated the last assessment undertaken of each participant. Changes in score constituted the dependent variables. The analyses were confined to the global impairment index (K-GAS) and the two main PTSD dimensions of re-experience of traumatic events and persistent avoidance of stimuli.

GAS impairment index: change from baseline to first-year follow-up. There was no significant main effect in relation to the two main treatment groups (individual v. group therapy). When analysis of covariance was undertaken in relation to the 'at home' and 'looked after' group, the effect size picture was not supported. Of the covariates, only the baseline score of the impairment index proved significant ($F=20.2$; significance of <0.001).

PTSD dimensions. For the re-experience of trauma dimension when the K-GAS impairment measure (log transformed) was used as a covariate, the findings are presented in Table 4. A significantly greater effect of individual therapy emerged in relation to change scores on this PTSD dimension at the first-year follow-up stage, baseline to second-year follow-up and baseline to exit from the study. Of the covariates, only the baseline score of impairment proved significant at second-year follow-up and at exit from the study. When the impairment index, as a covariate, was replaced by the initial score of the re-experience of trauma dimension (log transformed), and focusing on change from baseline, all the significant differences were reduced: baseline to first-year follow-up proved not significant; for baseline to second-year follow-up and exit from the study the F values and significance of F were 5.1 with $P<0.03$ and 4.82 with $P<0.032$, respectively. Re-experience of trauma (log transformed) at baseline as a covariate had F values and significance of F as 33.2 with $P<0.001$ and 40.6 with $P<0.001$, respectively. There were no significant interactions.

For the *persistent avoidance of stimuli* dimension, on analysis of covariance when the impairment index (log transformed) was a covariate, there was only one significant difference at the first-year follow-up

Table 4 Post-traumatic stress disorder: summary of analysis of covariance

	Time point	F	Significance of F	Covariate: impairment score at baseline	
				F	Significance of F
Re-experience of trauma	Baseline to first-year follow-up	4.3	<i>P</i> < 0.05	3.1	NS
	Baseline to second-year follow-up	7.5	<i>P</i> < 0.01	4.15	<i>P</i> < 0.05
	Baseline to exit from study	7.3	<i>P</i> < 0.01	14.5	<i>P</i> < 0.001
Persistent avoidance of stimuli	Baseline to first-year follow-up	5.5	<i>P</i> < 0.03	4.43	<i>P</i> < 0.05

NS, not significant.

Table 5 Change in impairment of overall functioning from baseline to first-year follow-up

Baseline	First-year follow-up			Total
	Major or serious impairment (K-GAS category 2–5)	Moderate impairment (K-GAS category 6)	Little impairment (K-GAS category 7–9)	
Major or serious impairment (K-GAS category 2–5)	13	10	13	36 (62)
Moderate impairment (K-GAS category 6)	3	5	10	19 (31)
Little impairment (K-GAS category 7–9)	0	0	4	4 (7)
Total	16 (28)	15 (26)	27 (47)	58

Pearson's $\chi^2=76.0$, d.f.=8, *P* < 0.001.

For the purposes of this analysis 'major' and 'serious' impairment categories have been combined. Values in parentheses are percentages. K-GAS, Kiddie Global Assessment Scale.

Table 6 Global assessment of functioning: shift in impairment categories across time

Category group	<i>n</i> (%) of children in group			
	Baseline	First-year follow-up	Second-year follow-up	At exit from study
A. Major impairment (K-GAS 2–4)	21 (30)	6 (10)	2 (4)	2 (3)
B. Serious impairment (K-GAS 5)	23 (32)	10 (17)	5 (9)	7 (11)
C. Moderate difficulty (K-GAS 6)	22 (31)	15 (26)	18 (33)	23 (35)
D. Limited impairment (K-GAS 7–9)	5 (7)	27 (47)	29 (54)	34 (51)
Sample size (<i>n</i>)	71	58	54	66 ¹

1. Number of subjects with data at baseline and second-year follow-up or exit.

Chi-squared test for categories A plus B, and C plus D: baseline to first-year follow-up $\chi^2=14.8$, *P* < 0.001, OR=4.3, 95% CI 2.0–9.2; baseline to second-year follow-up $\chi^2=29.9$, *P* < 0.001, OR=11.0, 95% CI 4.3–28.1.

K-GAS, Kiddie Global Assessment Scale.

(see Table 4). When the impairment index was replaced by the initial score on the persistent avoidance dimensions at baseline (log transformed), there was no significant difference.

Change irrespective of therapy

Univariate analyses

The K-GAS impairment data were split into three levels consisting of *major or serious impairment* (categories 2–5); *moderate impairment* (category 6); and, finally, some subjects with disorders or with multiple

symptoms but only *little impairment* – that is, 'usually functions well' (categories 7–9) (Table 5). At first-year follow-up there was a substantial shift away from major or 'serious' impairment towards 'moderate' or 'little' impairment: 62% to 28%. Of the 36 participants with 'major' or 'serious' impairment at baseline, 23 moved into categories of lesser degrees of impairment. Of the 18 children with 'moderate' impairment at baseline, this shift is compounded by movements in both directions: three moved to more serious degrees of impairment, and 10 moved to lesser degrees of

impairment. Finally, for the few children who were symptomatic but who were considered to have little impairment of functioning at baseline, there was no change. Notably, at baseline, 7% fell into the 'little impairment' category; at first-year follow-up; this figure was 47%. Table 6 illustrates the pattern of shift away from impairment over the course of time: 'major' or 'serious' impairment of 62% at baseline, to 27% at first-year follow-up, to 13% at second-year follow-up, to 14% at exit. In contrast, 'little' impairment increased from 7% at baseline to 47% at first-year follow-up, to 54% at second-year follow-up, and to 51% at exit.

Shift of psychiatric disorders over time. The psychopathological shift could be confined to the specific presence of psychiatric disorders at baseline (time 1) and at first-year follow-up (time 2); however, this would obscure the shifts that occur in both directions, namely disorder at time 1 and no disorder at time 2; no disorder at time 1 and disorder at time 2; and finally no disorder at time 1 and disorder at time 2 (Oates *et al*, 1994). Table 7 shows that 26 children (45%) had a general anxiety disorder at time 1, which reduced to 10 (17%) at time 2. However, the 17% is a combination of a

Table 7 General anxiety disorder: psychopathological disorder shifts from baseline to first-year follow-up

Baseline	First-year follow-up		Row total
	No disorder	Disorder	
No disorder	27 (84) (56)	5 (16) (50)	32 (55)
Disorder	21 (81) (44)	5 (19) (50)	26 (45)
Column total	48 (83)	10 (17)	58 (100)

Values in parentheses are percentages.

reduction in disorder in the children who had a disorder at baseline and new disorders in children who had no disorder at baseline.

Overall, the pattern consists of a systematic reduction in psychopathological symptoms from baseline to the first-year follow-up in all the listed disorders, with a small increase in psychopathological disorders in those who were free from such disorders at baseline. This pattern of shift is noted with all disorders and proved significant using McNemar's test (McNemar, 1949) for three of them: general anxiety ($P < 0.01$), depression ($P < 0.001$) and separation anxiety ($P < 0.001$). This is exemplified as follows in relation to these three disorders:

- Of the 26 participants with a general anxiety disorder at baseline, 21 no longer had this disorder at the first-year follow-up (a reduction of 81%); of the 32 without this disorder initially, five developed a disorder at the first-year follow-up (an increase of 16%) (Table 7).
- Of the 32 participants with depressive disorder at baseline, 22 were no longer depressed at the first-year follow-up (a reduction of 69%); of the 26 without this disorder, none developed a new disorder at the first-year follow-up (data not tabulated).
- Of the 36 participants with separation anxiety disorder at baseline, 18 no longer had this disorder at the first-year follow-up (a reduction of 50%); of the 22 without this disorder, two developed a disorder at the first-year follow-up – an increase of 9% (data not tabulated).
- Overall, the rates of disorder at the first-year follow-up were as follows:

general anxiety 17%, depression 17%, separation anxiety 23%, conduct 13% and simple phobias 10%.

Another way of studying change in psychiatric disorder is to note any changes in the mean comorbidity score over time. At baseline the mean comorbidity score was 2.59 (s.d.=2.03, 95% CI 2.10–3.07); at the first-year follow-up the mean score was 1.19 (s.d.=1.52, 95% CI 0.71–1.62), which is a reduction of 54% (or 0.7 s.d.); at the second-year follow-up the mean score was 0.92 (s.d.=1.1, 95% CI 0.61–1.23), which is a further reduction of 10%, making a 64% reduction over the 2 years (or 0.8 s.d.). Most of the reduction occurred over the first year.

Magnitude of change irrespective of treatment type

The intention here is to study the magnitude of change in the whole sample, irrespective of whether there are between-group differences. For these purposes, analysis of variance was employed on the four main measures – the K-GAS impairment measure and the three main dimensions of the PTSD. Data can be inferred from viewing Tables 2 and 3 and employing repeated measures analysis (at baseline and at the three subsequent time points). On all four dimensions, statistically significant changes were found: these were substantial on the K-GAS impairment index at first-year follow-up and increased further by the second-year follow-up. The reduction in mean impairment in standard deviations from baseline to the first-year follow-up was 1.16, at the second follow-up 1.43 and at exit 1.41, respectively. Similarly, on the three PTSD dimensions the changes were substantial at the first-year follow-up and this improvement was maintained at the second-year follow-up. However, the sexualised behaviour items (whether or not included in the PTSD algorithm) did not influence the picture materially, at either the first-year or the second-year follow-up.

DISCUSSION

Relation of CSA to mental health problems in children

Childhood sexual abuse is an event (or series of events) but it does not constitute a diagnosis or a disorder. However, the impact of the penetration, of threatened

and actual violence, and of the transgression of normal expected boundaries, leaves the abused child vulnerable. Indeed, there may already be degrees of vulnerability due to poor attachments, disruption of care and parental stresses, in which circumstances the CSA may act as a major risk factor leading to visible mental health problems, symptoms or disorders.

The extent of the psychiatric psychopathology before therapy in the cohort studied was considerable, including depression, separation anxiety and PTSD, and was often chronic (Trowell *et al*, 1999; further details available from the author upon request). Jones & Ramchandani (1999) highlighted the fact that about half the children in their survey who had been sexually abused experienced depression, PTSD or disturbed behaviour. Thus, the London sample was a rather more severely affected group than the range of samples studied so far and had substantial comorbidity and chronicity. On the basis of the above preconditions, it seemed unlikely that many of the presenting DSM disorders would improve spontaneously. In these circumstances, information about between-treatment effects following psychotherapy is of crucial importance.

Outcome

The univariate analyses examining between-treatment differences in change scores looked at effect sizes at specified criterion levels. On the K-GAS impairment measure, the effect sizes proved low at the first-year and second-year follow-ups and at exit from the study, and have been omitted from Table 2.

On the PTSD re-experience of traumatic events dimension, the effect size was 0.6 at the first-year follow-up, 0.79 at the second-year follow-up and 0.65 at exit; these were all in favour of individual therapy and they were all at the specified criterion level. Regarding persistent avoidance of stimuli, there was an effect size of 0.66 in favour of individual therapy at the first-year follow-up; this did not achieve the criterion level at the second-year follow-up but was 0.60 at exit. On analysis of covariance, there were three significant differences all in favour of individual therapy: these concerned re-experience of traumatic events (baseline to first-year follow-up; $P < 0.05$; baseline to second-year follow-up, $P < 0.01$; baseline to exit, $P < 0.01$). On the 'persistent avoidance of stimuli'

dimension there were two effect sizes above the specified criterion level, namely baseline to first-year follow-up and baseline to exit (0.66 and 0.60, respectively).

However, effect sizes do not adjust for relevant covariates, and when such adjustment was undertaken with analyses of covariance there emerged three significant differences on the 're-experience of trauma' dimension, all at least at the $P < 0.05$ level, all paralleling the effect sizes. For the 'persistent avoidance of stimuli' dimension there was only one significant main effect covering baseline to first-year follow-up. Such differences reduce when the impairment measure is replaced as a covariate by the baseline score of the outcome measure under scrutiny. There was no other significant difference. The above findings stand in contrast to the earlier accounts of outcome in the literature but are consistent with more recent conclusions from meta-analyses that individual therapy is more effective than group therapy (Weisz *et al*, 1995). In summary, the findings indicated that there were few statistically significant treatment differences, although the question of the small sample size and lack of power remains an issue.

Finally, one other main effect as represented by an effect size at the specified criterion level was identified, whether the child was 'looked after' or 'at home' on the K-GAS impairment measure. At the first-year follow-up it was in favour of the group that remained at home. This was not complemented by a significant difference on analysis of covariance. The above findings should be viewed against two major studies: first, the naturalistic study by Tebbutt *et al* (1997); second, the experimental study of Deblinger *et al* (1996).

Can 'naturalistic' treatment studies throw light on whether any 'overt' improvement identified is due to therapy or to 'maturation' or is it a regression to the mean effect? Numerous studies (not all scientifically sound) using a repeated-measures design suggest that therapy makes a contribution, as do those non-randomised designs where the treatment is clearly specified; however, Stevenson (1999) argues strongly that these do not constitute definitive statistical evidence of efficacy and further questions the validity of such findings. In contrast, Jones & Ramchandani (1999) also adopt an evidence-based approach but take the view that research should seek to support practice and not dictate it. They have, therefore, critically

examined the research available, gathering key findings that are relevant to professionals working in this field. For instance, they point out that treatments that involved only children who were symptomatic were particularly effective (Cohen & Mannarino, 1996; Deblinger *et al*, 1996). Another finding is that preschool children who drop out of therapy tend to be from families of lower socio-economic status (Cohen & Mannarino, 1997).

Naturalistic studies show a complex picture. They seem to support the view of Stevenson (1999). For instance, Tebbutt *et al* (1997), with the benefit of a 5-year follow-up of 68 sexually abused children, provided overall negative findings. Of the children who were sad or depressed at intake, 30 (44%) were improved at follow-up; 28 (41%) who were asymptomatic at intake had deteriorated. However, the conclusions drawn from this study about the absence of effects of treatment need to be viewed with caution as there is insufficient information about the quality and duration of any available treatment. Similarly, Oates *et al* (1994) reported a relative failure of 64 abused children aged 5–15 years to show significant improvement when reassessed 18 months later for depression and other psychopathological problems: 65% of the abused children had received therapy for an average of 9 months, the depression and behaviour scores of random community controls showed little change over time, and few of the abused children's scores moved towards a normal range. Nevertheless, 48% remained in a dysfunctional range for behaviour and 35% for depression. Improvement in the child's behaviour related to improvement in family functioning, but no relationship was found between therapy and outcome. The latter findings are in contrast to both the study by Deblinger *et al* (1996) and the London findings reported here; questions therefore arise about the quality and duration of treatment in the studies by Tebbutt *et al* (1997) and by Oates *et al* (1994). However, what cannot be discounted is the implication from the study by Tebbutt *et al* (1997) that sexually abused children with serious or chronic psychiatric disorders do not appear to respond to inadequate therapy (or to no therapy at all).

Experimental studies

The outcome study by Deblinger *et al* (1996, 1999) of cognitive-behavioural

therapy gives reasonable confidence of scientific soundness. The differential effects of child or non-offending mother participation were explored in a cognitive-behavioural intervention designed to treat PTSD and other emotional difficulties in school-aged sexually abused children. Those presenting for treatment were likely to have more overt and worrying symptoms complicated by comorbidity. Children and their families were assigned to one of four experimental conditions: these were child only; mother only; mother and child; and a standard community care situation, where parents were provided with information about where to seek therapy. The three active interventions were designed to target sexually abused children's PTSD symptoms, as well other emotional problems. Twenty-five children were allocated to each of the four experimental conditions. The therapy was conducted by mental health therapists trained in cognitive-behavioural interventions. Evaluations before and after treatment included standard measures of children's behavioural problems, anxiety, depression and PTSD symptoms, as well as measures of parenting practices. Three months after the inception of specific treatment, where there was active intervention, the mothers described significantly greater decreases in their children's externalising behaviour and an improvement in their own parenting skills; their children showed significantly greater decreases in their self-reported levels of depression. Further, there was a significantly greater reduction in PTSD symptomatology for children assigned to active treatment. An important relevant implication was a need to engage non-offending mothers in the treatment of their sexually abused children. In London, this policy was standard practice and complemented both the individual and group therapy.

The follow-up suggests that the improvements following treatment were stable across the 2-year follow-up period on the measures of psychopathological disorders that had significantly decreased at the end of therapy in the original study: externalising behaviour problems, depression and PTSD symptoms (Deblinger *et al*, 1999). However, questions remain about subsequent input such as booster sessions or other clinical interventions between the end of therapy and the 2-year follow-up; also, the rate of loss of data is high and a formula for correcting such substantial losses can give rise to distortions.

Nevertheless, the Deblinger *et al* study does support the notion that 'specific' therapy is far more helpful than non-specific community approaches. Whereas the Deblinger *et al* study demonstrates that active therapy is more helpful than 'inactive' approaches, the London study demonstrates a few significant between-group differences when comparing two active therapies. Thus, the Deblinger *et al* study and the London study provide evidence that demonstrable improvement with active therapy does not reduce over time. Further, Jones & Ramchandani (1999) made a contribution to this debate when reviewing psychological treatments: they indicate that, despite the view that any improvement noted may have been due to the passage of time, the results of some studies strongly suggest a treatment effect (Cohen & Mannarino, 1996; Deblinger *et al*, 1996).

Short-term therapy and long-term outcome

A crucial implication for outcome studies is that short-term therapy has to be viewed against both medium-term and longer-term outcome to ascertain whether treatment effects are maintained and whether new psychopathological problems emerge, balanced against selectivity in recruitment of the treatment sample (Kolvin *et al*, 1981, 1988). For example, the London cohort contained more seriously abused and seriously disturbed children with chronic disorders than are found in many other studies. Over 30% had been exposed to multiple perpetrators, compared with 7% in the study by Deblinger *et al* (1996). The Deblinger *et al* referral and inclusion criteria related to PTSD; not surprisingly, therefore, the psychiatric psychopathology at baseline was moderate. As the London inclusion criteria were that the child should be symptomatic, there were high rates of DSM-IV psychiatric disorders (Trowell *et al*, 1999); additionally, there is evidence of serious family dysfunction (further details available from the author upon request). These issues are discussed by Tebbutt *et al* (1997), who pointed out that most prospective outcome studies did not extend their second follow-up period beyond 18 months; furthermore, when there is a follow-up assessment (beyond 18 months) the attrition rates are usually high – for example 46% (Conte, 1987) – and only occasionally low – for example 20% (Oates *et al*, 1994), 19% (Tebbutt *et al*, 1997); and

in the current London study the rates were about 20% at first follow-up and under 10% at exit from the study. Such factors may compromise the generalisability of the findings.

Removal of children from home

Current practice is to remove children from an 'unsafe' abusive environment. The London research shows only a few outcome differences between children remaining in a prescribed 'safe' home environment and those removed that are in favour of the child remaining at home. Clinical practice indicates that there are both positive and negative effects of relocating some children. When such decisions are made, it is important that due weight is given to the disruptive effect of removal. What is remarkable in the London study is the equivalent outcome for children removed from home – they had been abused for longer, at a younger age of onset and by more abusers (further details available from the author upon request). It is impossible to predict what the outcome for them would have been if they had remained in the home situations in which abuse had occurred and where they would most probably have been vulnerable to some form of further abuse.

Change irrespective of treatment

Severity of impairment of global functioning and psychiatric disorder

Irrespective of treatment, there was a substantial and systematic reduction in both levels of impairment and extent of psychiatric disorder over the 2 years of the study. However, although there was a major reduction in impairment by the first-year follow-up, only about half of the children studied had moved into the category of 'limited' impairment or 'generally functioning pretty well'. The same was true for the psychiatric disorders: for instance, at the first-year follow-up, depression was substantially reduced, and by the second-year follow-up (although some children still had some depressive symptoms) none was considered to have a disorder. With general anxiety disorder there was a major reduction of disorder in participants who had presented with a disorder at the baseline assessment, but a minor movement towards the disorder in children who had showed little or no evidence of general anxiety

disorder at baseline. This phenomenon of psychopathological shifts in two directions in children who have been seriously sexually abused has only recently been studied and reported (Oates *et al*, 1994; Tebbutt *et al*, 1997; and the current study), and can easily be missed if the focus is to explore remission in only those subjects who present with a specific disorder at an earlier point.

Overall, the rates of major depressive disorder, general anxiety disorder and separation anxiety disorder at first-year and second-year follow-up were dramatically reduced. The pattern for PTSD was complicated (as reported above): there was a highly significant reduction in mean scores on the main dimensions by the first-year follow-up and this picture was confirmed at the second-year follow-up. It is evident that most of the change had occurred by the first-year follow-up.

There were few significant differences between the two forms of treatment and this may be attributed to the shared qualities of the interventions: sensitivity, understanding and concern, plus abuse-specific therapy (Jones & Ramchandani, 1999).

Dissociation

The findings showed that a small number of girls experienced an increase in symptomatology and impairment over the treatment period. This may be linked to the phenomenon of dissociation when individuals appear to be cut off from their traumatic experiences and where therapy allows them to process previous events. Short-term therapy may open up these experiences but be insufficient, and therapy of longer duration may be required.

What is known about spontaneous improvement?

Kendall-Tackett *et al* (1993), in their community-based study, reported that about a third of sexually abused children failed to exhibit overt psychological symptoms, or their symptoms might have already resolved by the time of initial assessment. About 55% of such children show substantial recovery from psychological symptoms over the subsequent 18 months, but 10–24% appear to become worse (Kendall-Tackett *et al*, 1993). Until recently, little information has been available about spontaneous remission in subjects in whom the psychopathology is more complex or severe, as in the current

study where children are referred for treatment.

Untreated control group

There is no rigorous study of the efficacy of psychotherapy in sexually abused girls; ideally, such a study would have a randomised design including an untreated control group, with all participants of school age, a sample size per group of 25 or more, with the therapy specified, and of appropriate duration with the subjects monitored for a year from intake into treatment, and the research would be published in a peer-reviewed journal (Finkelhor & Berliner, 1995; Stevenson, 1999). One of the reasons for the absence of such a study is that ethical considerations preclude the withholding of treatment from abused children for such a long time (Trowell *et al*, 1995). In drawing conclusions about 'spontaneous' improvement, reviewers need to consider (a) the severity of the abuse experiences, (b) the nature and complexity of associated family dysfunction, (c) the length of time since the last abuse episode, (d) whether the children were adequately protected after disclosure, and (e) the severity of psychopathological disorder at intake. Without such information it is difficult to draw general inferences that may aid clinicians.

Clinicians will be most interested in the evidence from the current study of the magnitude of changes in psychopathological disorders and impairment over time irrespective of the type of treatment given, but would appreciate explanations for these changes or even theoretical speculations. For example, Tebbutt *et al* (1997) asserted that their study sample consisted of more severely abused children than those of a similar age reported by Gomes-Schwartz *et al* (1990). It is argued that this is true, too, of the study reported here. This gives rise to an enigma: on the one hand, a high proportion of children in the London sample had severe and chronic disorders; according to Oates *et al* (1994) and Tebbutt *et al* (1997) such subjects do not improve over time, irrespective of treatment; furthermore, other research indicates that such disorders in sexually abused children do not improve unless treated (Deblinger *et al*, 1999). On the other hand, statistical experts, when viewing the London findings of changes over time irrespective of treatment, which are of considerable magnitude, are likely to see them either as representing spontaneous

improvement or as regression to the mean. However, this sample with its high levels of acute and chronic psychopathologies contains the very set of psychological problems that the two therapies have targeted, so clinicians deduce that therapy has contributed to the rapid decline of psychopathology. Without a control group, neither a statistical nor a clinical hypothesis can be fully or partially supported or rejected. On balance, none of the above explanations is fully convincing and this phenomenon of change merits further research.

Other relevant issues

Engagement and participation

The finding of Deblinger *et al* (1996) was that there was a need to engage non-offending mothers in the treatment of sexually abused children. Given the high levels of psychopathological problems in the children in the present study, and of parental and family dysfunction in the birth families (Rushton & Miles, 2000; further details available from the author upon request), there was surprisingly good engagement with the children and their carers, and comparatively low rates of drop-out. The majority of parents in the study were satisfied with the support they received, which probably was crucial in both the engagement and the maintenance of treatment. However, follow-up assessments proved more problematic.

Is major depression in sexually abused girls an intransigent disorder?

There is evidence that childhood major depressive disorder is not a transitory condition; it is a high-risk factor in parasuicide and may persist into adulthood (Kovacs, 1997). Other work emphasises this theme in relation to sexually abused children (Tebbutt *et al*, 1997). Two studies have found no change in symptoms of depressive disorders over time even with treatment (Oates *et al*, 1994; Tebbutt *et al*, 1997). In contrast, in the London study, irrespective of the specific treatment given, there is evidence of substantial reduction in rates of major depressive disorder, at the second-year follow-up. This suggests that, under certain conditions, major depression in sexually abused girls is not intransigent, despite the severity and chronicity of the psychopathological problems in the London

sample. These are matters of clinical importance and merit further research.

Are PTSD symptoms transient phenomena?

A crucial issue concerns the extensive presence of PTSD in various studies of psychological trauma. In adult subjects with exposure to psychological trauma (such as war trauma in veterans who sought psychiatric treatment 4–6 years afterwards), extensive PTSD comorbidity was diagnosed (Bleitch *et al*, 1997). Both lifetime (100%) and current (87%) PTSD was diagnosed, with major depressive disorder being the most prevalent comorbid diagnosis (95% lifetime, 50% current). This pattern does not differ greatly from that reported in childhood sexual abuse: PTSD is known to show a tendency to continuity across time, is often resistant to treatment, and has an impact on behaviour and social development; thus, children who develop PTSD may be at considerable risk for more prolonged dysfunction (McLeer *et al*, 1994). Of clinical importance is the response in favour of individual psychotherapy on one of the PTSD dimensions.

Limitations of the study

Design

The design consisted of a comparison of two active treatment groups. Two ethical committees declined the use of a no treatment or minimal treatment control group. However, some authorities may express uncertainty about whether any intervention works (Stevenson, 1999), and further argue that it is mandatory for all trials of psychotherapy in sexually abused subjects with psychiatric problems to include an untreated control group. This is a theoretical view, but the study by Deblinger *et al* (1999) has demonstrated that a control group is not always a feasible option, and also has provided some convincing evidence that specialised psychotherapy has much to commend it. This remains a contentious matter.

Sample size

The treatment groups were small, which resulted in reduced power in statistical analysis. The original intention was to recruit 90 subjects to therapy; however, eventually 81 children were assessed and, of these, 71 entered therapy. Although most participants had a sufficiency of therapy, engaging

them in follow-up did not prove easy and eventually only 58 children were assessed at the first-year follow-up. Thus, whereas on some measures (with the exception of PTSD dimensions), the two therapies appear equally effective, the low sample size has the drawback that absence of a statistical difference does not mean that such a difference does not exist, because a larger sample might well give rise to one.

Attrition

The attrition level was higher than hoped (18% by the end of the first year; 24% by the end of the second year); nevertheless, members of the research team were in contact with 66 participants (93%) at either the first-year or the second-year follow-up.

Masking

The therapies were so different that masking of subjects was impossible (Day & Altman, 2000). Masking of assessors was likely to be compromised because the children and their mothers often mentioned the specific therapy during assessment. This was accepted as an unavoidable limitation for the interview data (but not for the self-report inventories).

Level of dysfunction

This was a cohort with a high level of psychological disturbance and family dysfunction, and so the results can be generalised only to patients attending similar specialist child and family mental health clinics.

APPENDIX

Measurement issues

Further studies of the K-SADS have been undertaken with the aim of improving its structural and reliability qualities and its DSM-III-R or DSM-IV compatibility (American Psychiatric Association, 1987, 1994). One of these studies has addressed video-tape reliability, including DSM-III-R categorical agreement using kappa statistics (Ambrosini et al, 1989). The latter study addresses previous reliability concerns with regard to anxiety disorders and also behaviour disorders. Examples of the overall mean κ achieved for various disorders were as follows: major depression 0.80; overanxious disorder 0.85; separation anxiety 0.85; simple phobia 0.64; attention deficit 0.88. This pattern of reliability has been confirmed when administering the K-SADS to a clinical sample of subjects (G. Kolaitis & T. Korpa, personal communication, 2001).

For our clinical research purposes, attention was given to some of the K-SADS' structural weaknesses: for example, we have added the schedule of Orvaschel (1989) for assessing PTSD. For reliability,

we were guided by the principles established by Ambrosini et al (1989) in their research. For instance, the diagnostic categorisations were intended to cover the preceding 12 months, rather than only the length of the major complaint (if the latter persisted for less than 12 months).

Categorisation issues

A distinction was made between syndromes and disorders: syndromes had a relevant patterning of symptoms but not a sufficiency of impairment of psychological functioning, which is regarded by some as an 'almost *sine qua non* for the diagnosis of psychiatric disorder' (Thompson, 1989). Four of the five children with mild impairment ratings (K-GAS ratings of 7) were considered syndromic; the other child manifested diverse symptoms that did not conform to any main Axis I category. None of the other children included in therapy had K-GAS baseline ratings of 8 or 9.

Diagnostic issues

Diagnoses were assigned if the patient met the DSM criteria within that specified time frame. However, our study provides only global rather than diagnostic-specific impairment ratings (Kaufman et al, 1997). Finally, a *post hoc* review of all the K-SADS data was undertaken by two assessors (I.K. and H.S.) to provide consensus best-estimate DSM diagnoses. All the assessments (including that of PTSD) were undertaken by psychiatrists: the north London group provided 70% of the participating children and the south London group 30%.

Masking issues

At the baseline assessment the interviewers of both the parents and the participating children were masked as to treatment allocation; however, this was not possible at the follow-up assessments because the children often mentioned their therapy during the interviews. As Day & Altman (2000) state, in trials with alternative therapies 'full blinding is often impossible'; this limitation was appreciated. One way of minimising the potential bias was by reviewing the DSM diagnostic categories and also the coding of global impairment by the consensus checks alluded to above – the rationale being that any bias was more likely to occur at the pivotal point of making a clinical judgement based on the available data.

ACKNOWLEDGEMENTS

We thank Ildi Sherwood for administrative and secretarial help. I.K. was supported by a Leverhulme Emeritus Fellowship. The research was funded by the Department of Health and the Mental Health Foundation. The Tavistock Centre has provided additional support and resources.

REFERENCES

Ambrosini, P. J., Metz, C., Prabucki, K., et al (1989) Videotape reliability of the third revised edition of the K-SADS. *Journal of the American Academy of Child and Adolescent Psychiatry*, **28**, 723–728.

American Psychiatric Association (1987) *Diagnostic and Statistical Manual of Mental Disorders* (3rd edn revised) (DSM-III-R). Washington, DC: APA.

— (1994) *Diagnostic and Statistical Manual of Mental Disorders* (4th edn) (DSM-IV). Washington, DC: APA.

Begg, C., Cho, M., Eastwood, S., et al (1996) Improving the quality of reporting of randomised controlled trials. The CONSORT statement. *JAMA*, **276**, 637–639.

Beitchman, J. H., Zucker, K. J., Hood, J. E., et al (1991) A review of the short-term effects of child sexual abuse. *Child Abuse and Neglect*, **15**, 537–556.

Bleitch, A., Koslowsky, M., Dolev, A., et al (1997) Post-traumatic stress disorder and depression. An analysis of comorbidity. *British Journal of Psychiatry*, **170**, 479–482.

Brown, G. & Anderson, B. (1991) Psychiatric morbidity in adult inpatients with childhood histories of sexual and physical abuse. *American Journal of Psychiatry*, **148**, 55–61.

Casey, R. J. & Berman, J. S. (1985) The outcome of psychotherapy with children. *Psychological Bulletin*, **98**, 388–400.

Chambers, W. J., Puig-Antich, J., Hirsch, M., et al (1985) The assessment of affective disorders in children and adolescents by semi-structured interview. *Archives of General Psychiatry*, **42**, 696–702.

Cohen, J. (1960) A coefficient of agreement for nominal scales. *Educational Psychology Measurements*, **20**, 37–46.

— (1977) *Statistical Power Analyses for the Behavioural Sciences*. New York: Academic Press.

— (1992) A power primer. *Psychological Bulletin*, **112**, 155–159.

Cohen, J. A. & Mannarino, A. P. (1996) A treatment outcome study for sexually abused preschool children: initial findings. *Journal of the American Academy of Child and Adolescent Psychiatry*, **35**, 42–50.

— & — (1997) A treatment outcome study for sexually abused preschool children: outcome during a one-year follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry*, **36**, 1228–1235.

Conte, J. R. (1987) *The Effects of Sexual Abuse on Children: Results of a Research Project*. New York: New York Academy of Sciences.

Cotgrove, A. J. & Kolvin, I. (1996) The long term effects of child sexual abuse. *Hospital Update*, September, 401–406.

Day, S. J. & Altman, D. G. (2000) Blinding in clinical trials and other studies. *BMJ*, **321**, 504.

Deblinger, E., Lippmann, J. & Steer, R. (1996) Sexually abused children suffering post-traumatic stress symptoms: initial treatment outcome findings. *Child Maltreatment*, **1**, 310–321.

—, Steer, R. & Lippman, J. (1999) Two year follow-up study of cognitive behavioural therapy for sexually abused children suffering post-traumatic stress symptoms. *Child Abuse and Neglect*, **23**, 1371–1378.

Diggle, P. J. (1998) Dealing with missing values in longitudinal studies. In *Statistical Analysis of Medical Data* (eds B. S. Everitt & G. Dunn). London: Arnold.

Fine, S., Forth, A., Gillbert, M., et al (1991) Group therapy for adolescent depressive disorder: a comparison of social skills and therapeutic support. *Journal of the American Academy of Child and Adolescent Psychiatry*, **30**, 79–85.

Finkelhor, D. & Berliner, L. (1995) Research on the treatment of sexually abused children – a review and recommendations. *Journal of the American Academy of Child and Adolescent Psychiatry*, **34**, 1408–1423.

Gomes-Schwartz, B., Horowitz, J. M., Cardarelli, A. P., et al (1990) The aftermath of child sexual abuse: 18 months later. In *Child Sexual Abuse: The Initial Effects* (eds B. Gomes-Schwartz, J. Horowitz & A. Cardarelli), pp. 132–152. Newbury Park: Sage.

Guthrie, E. (2000) Psychotherapy for patients with complex disorders and chronic symptoms. The need for a new research paradigm. *British Journal of Psychiatry*, **177**, 131–137.

Hardy, G. E. & Shapiro, D. A. (1985) Therapists' verbal response modes in prescriptive vs. exploratory psychotherapy. *British Journal of Clinical Psychology*, **24**, 235–245.

Howard, K. I., Kopta, S. M., Krause, M. S., et al (1986) The dose–effect relationship in psychotherapy. *American Psychologist*, **41**, 159–164.

Jones, D. P. H. & Ramchandani, P. (1999) *Child Sexual Abuse. Informing Practice from Research*. Abingdon: Radcliffe Medical Press.

Kaufman, J., Birmaher, B., Brent, D., et al (1997) Schedule for affective disorders and schizophrenia for school-age children – present and lifetime version (K–SADS–PL): initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry*, **36**, 980–988.

Kazdin, A. (1998) *Research Design in Clinical Psychology* (3rd ed). Needham Heights, MA: Allyn & Bacon.

— (2000) *Psychotherapy for Children and Adolescents. Directions for Research and Practice*. Oxford: Oxford University Press.

Kendall-Tackett, K. A., Williams, L. M. & Finkelhor, D. (1993) Impact of sexual abuse on children: a review and synthesis of recent empirical studies. *Psychological Bulletin*, **113**, 164–180.

Kolvin, I., Garside, R. E., Nicol, A. R., et al (1981) *Help Starts Here: The Maladjusted Child in the Ordinary School*. London: Tavistock.

—, **Macmillan, A., Nicol, A. R., et al (1988)** Psychotherapy is effective. *Journal of the Royal Society of Medicine*, **81**, 261–266.

Kovacs, M. (1997) The Emanuel Miller Memorial Lecture 1994. Depressive disorders in childhood: an impressionistic landscape. *Journal of Child Psychology and Psychiatry*, **38**, 287–298.

Lanktree, C. & Briere, J. (1995) Outcome for therapy of sexually abused children: a repeated measure study. *Child Abuse and Neglect*, **19**, 329–334.

McLeer, S. V., Deblinger, E., Henry, D., et al (1992) Sexually abused children at high risk for post-traumatic stress disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, **31**, 875–879.

—, **Callaghan, M., Henry, D., et al (1994)** Psychiatric disorders in sexually abused children. *Journal of the American Academy of Child and Adolescent Psychiatry*, **33**, 313–319.

McNemar, Q. (1949) *Psychological Statistics*. New York: John Wiley & Sons.

Oates, R. K., O'Toole, B. J., Lynch, D. L., et al (1994) Stability and change in outcomes for sexually abused children. *Journal of the American Academy of Child and Adolescent Psychiatry*, **33**, 945–953.

Orvaschel, H. (1989) *Kiddies SADS–E Section. Designed to assess PTSD*. Philadelphia, PA: Medical College of Pennsylvania.

CLINICAL IMPLICATIONS

■ Focused individual and group therapy appear equally effective on measures such as the Global Assessment Scale, but for post-traumatic stress disorder individual therapy appears to have a greater impact.

■ Changes irrespective of treatment are of theoretical importance. The main changes are in the direction of health, with lesser shifts in the direction of psychopathology. Subjects who did not have a general disorder at baseline showed a move to psychopathology at the first- and second-year follow-ups.

■ It is important for sexually abused children to have a careful post-abuse screen assessment.

LIMITATIONS

■ The treatment groups were small, which resulted in reduced power in statistical analysis.

■ The cohort had a high level of psychological disturbance and family dysfunction, so the results can be generalised only to subjects attending specialist health clinics.

■ In the absence of a control group, the changes cannot be definitely attributed to the psychotherapeutic treatments.

J. TROWELL, FRCPsych, I. KOLVIN, FRCPsych, The Tavistock and Portman NHS Trust; T. WEERAMANTHRI, MRCPsych, South London and Maudsley NHS Trust; H. SADOWSKI (deceased), MRCPsych, Royal Free Hospital and Tavistock and Portman NHS Trust; M. BERELOWITZ, FRCPsych, Royal Free Hospital; D. GLASSER, FRCPsych, Great Ormond Street Hospital for Sick Children; I. LEITCH, MBCS, Consultant Statistician, London

Correspondence: Dr J. Trowell, The Tavistock and Portman NHS Trust, Child & Family Department, 120 Belsize Lane, London NW3 5BA, UK

(First received 7 March 2000, final revision 8 August 2001, accepted 1 October 2001)

Rushton, A. & Miles, G. (2000) Study of support service for current carers of sexually abused girls. *Journal of Clinical Child Psychology and Psychiatry*, **5**, 411–426.

Shaffer, D., Gould, M. S., Grasic, J., et al (1983) A children's global assessment scale (C–GAS). *Archives of General Psychiatry*, **40**, 1228–1231.

Smith, M. L. & Glass, G. V. (1977) Meta-analysis of psychotherapy outcome studies. *American Psychologist*, **32**, 752–760.

Stevenson, J. (1999) The treatment of the long-term sequelae of child abuse. *Journal of Child Psychology and Psychiatry*, **40**, 89–111.

Tebbutt, J., Swanston, H., Oates, R., et al (1997) Five years after child sexual abuse: persisting dysfunction and problems of prediction. *Journal of the American Academy of Child and Adolescent Psychiatry*, **36**, 330–339.

Thompson, C. (1989) *The Instruments of Psychiatric Research*. Chichester: John Wiley & Sons.

Trowell, J., Berelowitz, M. & Kolvin, I. (1995) Design and methodological issues in setting up a psychotherapy outcome study with girls who have been sexually

abused. In *Research Foundations for Psychotherapy Practice* (eds M. Aveline & D. Shapiro). Chichester: John Wiley & Sons.

— & **Kolvin, I. (1999)** Lessons for a psychotherapy outcome study with sexually abused girls. *Clinical Child Psychology and Psychiatry*, **4**, 79–89.

—, **Ugarte, B., Kolvin, I., et al (1999)** Behavioural psychopathology of child sexual abuse in schoolgirls referred to a tertiary centre: a North London study. *European Journal of Child and Adolescent Psychiatry*, **8**, 107–116.

Weisz, J. R., Weiss, B., Han, S. S., et al (1995) Effects of psychotherapy with children and adolescents revisited: a meta-analysis of treatment outcome studies. *Psychological Bulletin*, **117**, 450–468.

Yates, A. (1982) Children eroticized by incest. *American Journal of Psychiatry*, **139**, 482–485.

Zanarini, M., Gunderson, J. G. & Marino, M. F. (1989) Childhood experiences of borderline patients. *Comprehensive Psychiatry*, **30**, 18–25.