Therapist effects in routine psychotherapy practice: An account from chronic fatigue syndrome
Matteo Cella, Daniel Stahl, Silje Endresen Reme; Trudie Chalder

King's College, Department of Psychological Medicine, Institute of Psychiatry, London, UK
Department of Clinical, Educational and Health Psychology, University College, London, UK
King's College, Department of Biostatistics, Institute of Psychiatry, London, UK
University of Bergen, Department of Health Promotion and Education, Faculty of Psychology, Bergen, Norway

First published on: 24 January 2011

To cite this Article Cella, Matteo, Stahl, Daniel, Reme, Silje Endresen and Chalder, Trudie (2011) 'Therapist effects in routine psychotherapy practice: An account from chronic fatigue syndrome', Psychotherapy Research, 21:2, 168 — 178, First published on: 24 January 2011 (iFirst)
To link to this Article: DOI: 10.1080/10503307.2010.535571
URL: http://dx.doi.org/10.1080/10503307.2010.535571

Please scroll down for article
Therapist effects in routine psychotherapy practice: An account from chronic fatigue syndrome

MATTEO CELLA,1 DANIEL STAHL,2 SILJE ENDRESEN REME,3 & TRUDIE CHALDER1

1King’s College, Department of Psychological Medicine, Institute of Psychiatry, London UK, and Department of Clinical, Educational and Health Psychology, University College London, UK; 2King’s College, Department of Biostatistics, Institute of Psychiatry, London, UK & 3University of Bergen, Department of Health Promotion and Education, Faculty of Psychology, Bergen, Norway

(Received 23 June 2010; revised 20 September 2010; accepted 20 October 2010)

Abstract
The effect of therapists in psychotherapy is a much debated topic, with a number of studies showing therapist variance being large while other studies show little or no variability in outcomes due to therapists. The aim of this study was to investigate therapist effects in a well-defined sample of patients and therapists from an outpatient service which specializes in providing cognitive behaviour therapy (CBT) for patients with chronic fatigue syndrome (CFS). Therapy was provided in a highly specialized clinical setting for CFS and was delivered by qualified CBT therapists with at least 2 years experience with this client group. Three hundred and seventy-four patients with CFS and 12 cognitive behavioural psychotherapists took part. Therapist effects on the primary outcomes of fatigue and disability were investigated with multilevel random effects models and variance component analysis. Different models were computed and compared. Results showed a reduction in fatigue and disability scores after therapy. Variance explained by therapists, when demographic covariates were accounted for, was 0% for fatigue and under 2% for disability. A number of important factors may have played a significant role in minimizing therapist effects in our study. These are: specialist setting, single centre, patients with the same primary diagnosis, therapists of the same orientation and training, shared environment and supervision. Future studies may stress the importance of these factors in the investigation of the therapist effect in psychotherapy.

Keywords: cognitive behaviour therapy; outcome research; statistical methodology; therapist effects; multilevel analyses

Introduction
A substantial improvement in the evaluation of therapeutic interventions has been the consideration of the role of the service provider. The provider's influence on the effectiveness of an intervention has been perceived variably from different disciplines. Drug studies have tended not to stress the role of intervention providers in influencing therapeutic outcomes due to a more rigid application of double-blind design where both providers and service users are unaware of whether the treatment administered was active or not (Shapiro & Shapiro, 1997). In addition, the administration of interventions in drug studies requires generally a circumscribed set of skills with minimal or little interaction with the intervention provider. On the other hand investigations on treatment effectiveness in the fields of education and psychotherapy are believed to carry a stronger provider effect (e.g. Martindale, 1978). In these areas it is in fact difficult to achieve a double-blind design and contact between provider and service users is generally extensive and the source of intervention. Therefore ignoring the role of the provider in these areas may yield significant errors in estimating intervention effectiveness. In support of this notion a considerable wealth of research in these areas has shown that significant proportions of intervention variability can be accounted for by the provider (Huppert et al., 2001; Kim, Wampold, & Bolt, 2006; Lafferty, Beutler, & Crago, 1989). Also, as noticed by Kim et al. (2006), unless the provider effect is non-existent, which is often not the case, observations within the same provider are dependent. Not accounting for dependent observations in statistical analysis can result in the inflation of the effect size for treatment differences and increase the
chances of type I error (Crits-Christoph & Mintz, 1991; Kirk, 1995).

The interest in the effect of the provider in psychotherapy is not novel but although several studies have been conducted in this area, results are far from conclusive. Studies have shown that the therapist effect can range from 0% to almost 50% (Crits-Christoph & Mintz, 1991). In considering the causes of such high variability, various factors have been advanced as important in predicting the therapist effects; these factors can be tentatively grouped in three clusters: therapy type, patients diagnosis and analysis strategy.

In relation to therapy type, well-controlled and manualized therapies have been shown to produce smaller therapist effects (Crits-Christoph & Mintz, 1991), and therapist training and experience have been shown to reduce variability of outcomes, with inexperienced therapists associated with high variability (Huppert et al., 2001). Patient reporting of initial symptom severity and symptom heterogeneity is also thought to play a crucial role in accounting for therapist effects. Nevertheless, the large majority of the studies in this area have not properly accounted for potential symptom severity differences between patients at intake (e.g. Elkin, Falconner, Martinovich, & Mahoney, 2006).

Most relevant, perhaps, is the issue of the analysis strategy adopted. Although multilevel modelling has become the standard for therapist effect analysis (e.g. inputting therapist as random factor) several methodological issues are thought to play an important role in deflating the “real” therapist effects. These are: omitting from the analysis therapists at the tails of the distribution as outliers, inflating patient error variance by using the last score available for longitudinal trajectories, or choosing a statistical model that makes unreasonable assumptions about the phenomena studied, such as anchoring the intercept and thereby assuming that intake scores have no variance, unlike subsequent scores (Wampold & Bolt, 2006).

Further, the setting in which the study of therapist effects is conducted may exert an effect. More institutionalized settings may allow less flexibility from a psychotherapeutic point of view and this may clash with some therapists’ personal psychotherapy style. Similarly the therapeutic style predominant in the service may influence therapists with a different background. Further the homogeneity versus heterogeneity of clients treated could have a differential impact on therapists, with some improving their performance through experience with the specific client population and others reducing their effectiveness because of lack of stimulation. Different settings and study design (e.g. randomized controlled trials versus routine clinical practice) are also likely to have an effect on psychotherapy outcomes (Barkham et al., 2008). Although it has been suggested that randomized controlled studies should produce fewer therapist effects compared to routine clinical practice, this claim is not yet supported by empirical data in CFS (Quarmby, Rimes, Deale, Wessely, & Chalder, 2007). Besides, studies exploring therapist effects in routine clinical settings are few, even though patients are mostly treated outside clinical trial protocols. It seems therefore more important to establish the therapists’ contribution to therapy variance in naturalistic clinical environments rather than in research settings.

Studies using multilevel modelling (considering therapists as a random factor and assuming that treatment effectiveness varies randomly across therapists) hold the potential to generalize the findings to the entire therapist population. However, this level of generalization may be problematic to achieve for psychotherapy as a whole. Therapists with dissimilar therapeutic orientations may conceptualize disorders differently and apply theoretical formulations which reflect their orientation framework. These differences may contribute to a substantial variability in therapists and complicate the generalizability of the therapist effect findings. Therefore investigating the therapist effects in more homogeneous groups of patients and therapists can inform us about the characteristics that may contribute to modulating it.

Chronic fatigue syndrome (CFS) is a condition characterized by chronic disabling fatigue not better explained by an alternative diagnosis (Sharpe et al., 1994). As many as half the patients with CFS are unemployed (Lloyd & Pender, 1992) and present with various associated disorders. Depression and anxiety are the most common associated disorders, with a prevalence rising above 50% in primary care setting (Leone, 2010; Wessely, Chalder, Hirsch, Wallace, & Wright, 1996). The prognosis is poor: in primary care only a third improve in the first year, and of those referred to secondary care less than 10% return to pre-morbid functioning (Cairns & Hotopf, 2004).

There is now some evidence that specific treatments can improve these poor outcomes. A systematic review found that both cognitive behavioural psychotherapy (CBT) and graded exercise therapy (GET) are the most promising treatments for CFS in secondary care (Chambers, Bagnall, Hempell & Forbes, 2006). A Cochrane review recently confirmed that around 40% of patients with CFS report improvements in fatigue and physical functioning if they receive CBT (Price, Mitchell, Tidy & Hunot, 2008). In a non-randomized study similar outcomes were found in routine clinical practice (Quarmby et al., 2007).
The cognitive behavioural approach involves enabling individuals to develop a consistent approach to activity, gradually increase activity, develop healthy sleep patterns and identify and challenge unhelpful cognitions with the primary aim of improving fatigue and physical functioning (Sharpe & Chalder, 1994). In CFS, treatment is usually initially directed at behaviour change, with cognitive strategies being introduced once a consistent approach to activity and engagement has been established.

CBT for CFS is based on a cognitive behavioural model of the disorder. This model postulates that unhelpful catastrophic interpretations of symptoms and excessive focus on symptoms are central in driving disability and symptom severity (Chalder, Deale & Wessely, 1999; Deale, Chalder & Wessely, 1998). These cognitive responses are also associated with behavioural patterns which contribute to outcome, including avoidance of activity together with excessive rest, and all-or-nothing behaviour; a pattern of pushing too hard or being over-active when feeling well and then needing to rest up or do very little for prolonged periods. This model was originally influenced by theories of chronic pain and in particular the work of Philips (Philips, 1987).

The current study aims to explore the therapist effects in a group of patients fulfilling the diagnostic criteria for CFS referred to a specialized outpatient clinic. By minimizing the diagnostic heterogeneity (i.e. all CFS patients), the intervention type (i.e. all CBT) and provider variability (i.e. therapists with similar training and supervision), we anticipated that the therapist effects would be minimal or non-existent.

**Method**

For this study a subset of routinely collected assessment measures extracted from the Chronic Fatigue Research and Treatment Unit database was used. The original database comprised measures of both physical and psychological symptoms of patients with CFS treated within the unit across four time points: pre-therapy, post-therapy and two follow-up points, respectively after 6 and 12 months from therapy completion. One thousand two hundred and thirty-three entries collected from 1999 to 2007 were screened for inclusion in this study. Patients were excluded for the purpose of the current investigation if a full course of therapy was not completed, follow-up sessions were not attended, any of the primary outcome measures were not completed at any time point, if patients received CBT from a trainee psychotherapist under supervision or medical doctor with less than 2 years of experience in CBT, or if patients received other forms of therapy (e.g. graded exercise therapy).

The dataset used to investigate the therapist effects was created including patients presenting all the primary outcome measures (i.e. Fatigue Scale and WSAS) at the four assessment points considered. The choice of using only the unit primary outcomes to study therapist effects was dependent on the consistency with which these measures were used in the CFS unit in the time considered. Preliminary analyses did not detect differences in demographic variables (i.e. age and gender) between patients included and excluded from the current study. Also the proportion of missing data per number of patients treated did not differ significantly between therapists.

**Measures**

**Chalder Fatigue Scale**

This is an 11-item scale measuring physical and mental fatigue. Four response options are available for each item, ranging from “less than usual” to “much more than usual.” The scale total score can range from 0 to 33. The fatigue scale has been used extensively in research and has good psychometric properties (Cella & Chalder, 2010; Chalder et al., 1993).

**Work and Social Adjustment Scale**

This 5-item scale measures the degree to which the person’s fatigue affects their ability to work, engage in household management and participate in social and private leisure activities and relationships. The total score ranged from 0 to 40. Two different studies found the scale to be both a reliable and valid measure of impaired functioning (Mundt, Marks, Shear, & Greist, 2002). It was also found to be sensitive to patient differences in disorder severity and treatment-related change (Mundt et al., 2002).

**Participants**

**CFS Patients**

Three hundred and seventy-four patients met the inclusion criteria for this study. The patients included were initially referred from primary care to the CFS clinic, a specialist CFS clinic in the UK. All referrals were initially assessed by a consultant psychiatrist or senior therapist who confirmed the diagnosis of CFS. Inclusion criteria for the study were: (1) minimum age of 18 at intake; (2) a diagnosis of CFS according to the Oxford criteria.
Therapists and Treatment

Therapists were either clinical psychologists (three) or nurses with specialist CBT training (nine). All the nurses successfully completed an 18-month postgraduate diploma in cognitive-behavioural psychotherapy. All therapists were regularly supervised and had at least 2 years experience in delivering CBT. There were 10 female and two male therapists, with experience in CBT ranging from 2 to 17 years at the beginning of the study and experience with CFS clients ranged from 1 to 15 years. The number of CBT sessions ranged from 10 to 15 and were all delivered on an individual basis (i.e. one to one). Therapy was manualized and therapists were encouraged to structure each session and the entire intervention as outlined in the manual. Supervision ensured correct therapist adherence to the protocol used. The CBT offered included encouraging the patient to establish a consistent approach to activity and then building up activity gradually, the establishment of a sleep routine and addressing unhelpful cognitions related to fear of engaging in activity, perfectionism and high standards.

The average number of patients per therapist was 31.2 (SD = 24.4) with a range between eight and 122.

Analysis

Random effects models and variance component analysis were used to describe therapists’ variability and primary outcome score changes over time (Snijders & Bosker, 1999; Wampold & Brown, 2005). In the analyses patients were nested within therapists and repeated observations over time were nested within patients. Both therapists and patients were treated as random factors as assumed to represent a randomly selected sample of their respective populations. Using patients and therapists as random factors allows generalization of the current results to the respective underlying populations.

The possible effect of time in influencing primary outcomes was accounted for in the model as a continuous fixed effect. Preliminary analyses showed a non-linear relationship between change in primary outcomes and time; therefore time was included as a quadratic polynomial term in all the analysis. Age and gender were also included as fixed effect covariates. It was not possible to include the therapists’ characteristics due to the small number of therapists. The possible effect of time, as a linear or quadratic term, over primary outcome changes across patients (i.e. random slope) was also assessed and included in the analysis as a random effect. Model comparison showed time was better accounted for as a linear term. The small number of therapists and the risk of over-fitting did not allow for assessing the effect of time across therapists.

As time was allowed to vary across patients (i.e. random slope) variance explained by therapists and patients was relative to each time point. In order to control for this effect we performed an additional set of analysis to estimate variance components at post-treatment, 6 and 12 months follow-up. This analysis also allowed controlling for levels of fatigue and disability scores at intake (i.e. pre-treatment). Three different models were estimated for each of the primary outcomes:

(i) an unconditional model with a random effect of therapist only;
(ii) a conditional model controlling for initial level of severity before treatment by including baseline levels as a covariate;
(iii) a conditional model which included gender and age in addition to baseline severity measure.

The small number of therapists did not allow the slope of patient baseline scores to vary across therapists. Preliminary analysis showed that, when allowed to vary across therapists, including fatigue and disability scores as random slopes did not improve the model fit. Consequently only random intercept models are presented. Variance components were also estimated for the 6 month and 1 year follow-up.

Akaike’s Information Criterion (AIC) and Bayesian information criteria (BIC) were used to assess model fitness. Model selection was conducted according to
Burnham and Anderson’s (2002) recommendations and privileged parsimonious models that best explain the data with a minimum number of estimated parameters as identified with low AIC/BIC values. Notably, differences in AIC/BIC of less than 2 are regarded as negligible.

**Intraclass Correlation (ICC) and Model Fitting**

Proportions of therapists’ explained variance were estimated with intraclass correlation (ICC), defined as the ratio of the variance attributable to the therapists to the total variance (error variance + variance attributable to patients). In the simple random intercept model used the ICC ($\rho$) for therapists is:

$$\rho_{\text{therapist}} = \frac{s_{\text{therapist}}^2}{s_{\text{therapist}}^2 + s_{\text{error}}^2}$$

with

$$s_{\text{therapist}}^2 = \text{variance attributable to therapist}$$

$$s_{\text{error}}^2 = \text{error variance}$$

ICC values are a measure of the degree of independence of fatigue and disability scores nested by therapists. The ICC values ranges from 0 (complete independence) to 1 (complete dependence). Therefore an ICC score of 0 would suggest that the therapists are irrelevant to understanding the differences in symptom changes among patients, whereas an ICC score approaching 1 would suggest that individual differences in therapists are critical in understanding changes in patients’ symptoms.

Restricted maximum likelihood estimations were used for model fitting estimation. Model selection was based on information criteria (i.e. AIC and BIC, Burnham & Anderson, 2002) and log likelihood ratio tests, for nested models only. The main aim of studies estimating the variance components (such as reliability studies) is not testing the null hypothesis (i.e. the population ICC = 0) but an accurate estimation of the ICC (de Jong, Moerbeek, & van der Leeden, 2010). In addition to parameter estimates we therefore present 95% credible intervals around the estimate of the ICCs as measures of precision. A credible interval is the Bayesian analogue of a confidence interval. A 95% credible interval has a 95% probability of containing the true parameter value for the ICC. We calculated Bayesian credible intervals instead of relying on distributional approximations of standard frequency estimations of confidence intervals which are based on large sample sizes because of the small number of therapists and the expected small values of the ICC (Gelman et al., 2004). The Bayesian approach allows us to make exact inferences based on the posterior distribution for any sample size. Non-informative and vague priors on the parameters of the models as described in Bates and Maechler (2010) were used. These priors have little influence on inference. The credibility intervals were estimated by Markov chain Monte Carlo methods using 10,000 iterations.

Random effect modelling analysis assumptions were assessed by visual inspections of quantile-quantile plots of predicted random effects and residuals and scatter plots of predicted values versus residuals (Everitt & Hothorn, 2009). Analyses were performed in R 2.10 (R Development Core Team, 2009) using the lme4 library (Bates, 2009; Bates and Maechler 2010).

As patients were not randomly assigned to therapists, the potential difference on baseline variables was assessed. Both demographic characteristics and symptoms at baseline were not different between patients nested by therapists at pre-treatment.

**Results**

**Patients’ Characteristics**

Of the 374 patients considered for this study 66.3% were female. The mean age of the sample was 38.3 (SD10.7). Mean fatigue score at intake was 23.9 (SD6.6) and mean disability score was 25.4 (SD8.6).

**Change Over Time**

Figure 1(a) shows fatigue trends for patients grouped by therapist. A random effects model, with therapist and patient as random factors and time as a quadratic polynomial term, showed a significant non-linear reduction in fatigue scores across time. A small but significant increase in fatigue scores was observed around the first follow-up period. All patients, regardless of therapist nesting, showed the highest decrease in fatigue scores between baseline and end of treatment. Of all the different models attempted, the model with the best fit allowed time to vary across patients and included age and gender as covariates. In this model the estimated variance attributed to therapists was equal to zero, and age and gender did not significantly affect the outcome scores ($p = 0.11$ and $p = 0.29$ respectively; see Table I(a).

Figure 1(b) shows disability trends for patients grouped by therapist. A model similar to the one used for fatigue showed a non-linear reduction of disability scores across time. Again decrease was more evident between baseline and end of treatment compared to the follow-up assessments. In this case
only irrelevant amounts of variance could be attributed to therapists. A positive relationship between age and WSAS score change \((p = 0.013)\) was also detected (see Table 1b).

**Models of Variance Components for Fatigue**

According to Wampold and Brown (2005) and as specified in the analysis section, we performed three random factor models to estimate therapist’s variance. The first model performed was unconditional, with therapist as a random effect and fatigue scores at post-treatment as dependent variable. Results showed that the estimate of the therapist’s variance and therefore the intraclass correlation was 0. The second model was a conditional model on the baseline fatigue scores. Similarly to the unconditional model the conditional model showed an estimate therapists’ variance of 0. There was a significant positive relationship between baseline and post-treatment fatigue scores.

A third, conditional model was also performed including age and gender as covariates. Age showed a significant positive effect on fatigue scores but the ICC of therapist remained 0. Similar results were obtained using the two follow-up assessments. The ICC for therapist was virtually 0 on both occasions. The lower limit of the 95% credible intervals for the ICs was 0 for all models and the upper limits ranged between 2.8% and 7.6% (for full results see Table II).

**Models of Variance Components for WSAS**

The relevance of therapists on disability score was investigated with a similar modelling strategy to fatigue. The first model performed was unconditional, with therapist as a random effect and disability scores at post-treatment as dependent variable. Estimates of therapist’s variance, and therefore the intraclass correlation, was virtually 0.

The second model was conditional on the baseline disability scores. This model showed an estimate therapists’ variance of 0.9 with an ICC score of 0.018. A positive relationship between baseline and post-treatment WSAS scores was also found. A third, unconditional model including age and gender as covariates was performed. Adding gender and age as covariates did not alter the variance explained by therapists compared to the previous model (ICC = 0.017). Age was found to have a significant positive effect on disability scores at post-treatment and at 6 month and 1 year follow-up assessments. Gender was not found to influence WSAS scores. The ICC for therapist was less than 0.017 in every model (for full result see Table II). The lower limit of the 95% credible intervals was for all models 0 and the upper limits ranged between 5.8% and 9.5% (for full results see Table II).

**Discussion**

This report aimed to explore the effect of therapists on psychotherapy outcomes in a specialist service delivering CBT for CFS patients. We investigated this research question with different multilevel modelling strategies to account for different factors involved in the change of the primary outcomes. The results converge in suggesting that therapists have no influence in affecting the clinical outcomes of patients who have received CBT. This finding
Table I. Fatigue (a) and Disability (b) Estimates Obtained with Random Effect Analysis

<table>
<thead>
<tr>
<th></th>
<th>Fatigue (a)</th>
<th></th>
<th>Disability (b)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Random effect</td>
<td>Covariance</td>
<td>Random effect</td>
<td>Covariance</td>
</tr>
<tr>
<td></td>
<td>Variance (SE)</td>
<td>Covariance</td>
<td>Variance (SE)</td>
<td>Covariance</td>
</tr>
<tr>
<td>Patient</td>
<td>19.62 (4.43)</td>
<td>0.2</td>
<td>62.97 (7.94)</td>
<td>-5.88</td>
</tr>
<tr>
<td>Time</td>
<td>2.58 (1.61)</td>
<td>-0.2</td>
<td>5.78 (2.40)</td>
<td>-0.2</td>
</tr>
<tr>
<td>Therapist</td>
<td>0 (0)</td>
<td>0.33 (0.57)</td>
<td>0.92 (0.91)</td>
<td>0.01</td>
</tr>
<tr>
<td>Residual</td>
<td>26.69 (5.12)</td>
<td>20.68 (4.55)</td>
<td>0.01 (0.04)</td>
<td>0.01</td>
</tr>
<tr>
<td>Fixed effects</td>
<td>Estimate (SE)</td>
<td>30.88 (1.38)</td>
<td>27.76 (1.75)</td>
<td>1.04 (0.12)</td>
</tr>
<tr>
<td>Intercept</td>
<td>30.88 (1.38)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sex (Male =1)</td>
<td>-0.71 (0.67)</td>
<td>0.29</td>
<td>0.92 (0.91)</td>
<td>0.31</td>
</tr>
<tr>
<td>Age (years)</td>
<td>0.05 (0.03)</td>
<td>0.11</td>
<td>0.01 (0.04)</td>
<td>0.01</td>
</tr>
<tr>
<td>Time²</td>
<td>1.63 (0.13)</td>
<td>&lt;0.001</td>
<td>1.04 (0.12)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

A small number of recent investigations and theoretical contributions have started to indicate that the variability observed in therapist effects may be due to a set of under-considered variables, namely: clients’ diagnostic heterogeneity, therapists’ orientation, supervision, single vs multi-centre trials, and study setting and design (Crits-Christoph & Mintz, 1991; Elkin et al., 2006). Diagnostic heterogeneity and the role of comorbidity are factors that previous studies have addressed poorly by either failing to characterize the client group (e.g. Anderson, Ogles, Patterson, Lambert, & Vermeersch, 2009) or by investigating highly heterogeneous groups of clients with mixed and somehow different disorders (e.g. Dinger, Strack, Leichsenring, Wilmers, & Schauenburg, 2008; Elkin et al., 2006). In the current study, only patients fulfilling a diagnosis of CFS according to the Oxford criteria were included (Sharpe et al., 1991). Although we acknowledge that a diagnosis of CFS does not exclude per se the presence of comorbid features such as depression and anxiety, these were not primary complaints in our client group and they were all referred primarily for CFS and associated complaints.

Patient symptom heterogeneity may hence be an important and under-considered factor in the study of therapist effects since patients with broadly similar symptoms may reduce variability within the same therapist. Similar patients may also influence the therapist effects by creating a more homogeneous and shared environment for the therapists. Indeed, the therapists included in the current study, although having experience with different client groups, conduct therapy principally with CFS patients. The selected nature of the patients coupled with the specific training in the area and the shared nature of the supervision may be critical factors contributing to the small therapist effects observed.

A number of aspects need to be considered in interpreting the current results in the context of similar previous studies. Many of the previous studies on therapist effect were multi-centre studies (e.g. Luborsky et al., 1997). In our study, all patients recruited were diagnosed and treated in a single...
Table II. Parameter Estimates for Variance Components Analyses of Fatigue and Disability at Post-Treatment (Post), 6 Months Follow-Up (6mts) and 1 Year Follow-Up (1year)

<table>
<thead>
<tr>
<th>Time</th>
<th>Model</th>
<th>AIC</th>
<th>BIC</th>
<th>Therapist variance (SE)</th>
<th>Variance residual (SE)</th>
<th>ICC% (95% CI)</th>
<th>Intercept (SE)</th>
<th>p</th>
<th>Baseline</th>
<th>p</th>
<th>Age (SE)</th>
<th>p</th>
<th>Gender (SE)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatigue</td>
<td>Post</td>
<td>M1</td>
<td>2669</td>
<td>2681</td>
<td>0 (0)</td>
<td>72.71 (8.53)</td>
<td>0.000</td>
<td>0 (0-5.6)</td>
<td>17.6 (0.44)</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M2</td>
<td>2588</td>
<td>2604</td>
<td>0 (0)</td>
<td>59.94 (7.74)</td>
<td>0.000</td>
<td>0 (0)</td>
<td>4.1 (1.56)</td>
<td>** 0.56 (0.06)</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M3</td>
<td>2573</td>
<td>2597</td>
<td>0 (0)</td>
<td>60.06 (7.75)</td>
<td>0.000</td>
<td>0 (0-5.1)</td>
<td>2.13 (2.14)</td>
<td>p = 0.32</td>
<td>0.56 (0.06)</td>
<td>***</td>
<td>0.05 (0.06)</td>
<td>p = 0.21</td>
</tr>
<tr>
<td>6mts</td>
<td>M1</td>
<td>2670</td>
<td>2682</td>
<td>1.1E-07 (0.0003)</td>
<td>74.3 (8.62)</td>
<td>1.5E-9</td>
<td>0.000 (0-7.6)</td>
<td>16.62 (0.44)</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M2</td>
<td>2595</td>
<td>2611</td>
<td>1.3E-10 (1.1E-05)</td>
<td>62.21 (7.89)</td>
<td>2.1E-12</td>
<td>0.000 (0-5.9)</td>
<td>3.56 (1.59)</td>
<td>* 0.54 (0.06)</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M3</td>
<td>2574</td>
<td>2598</td>
<td>0 (0)</td>
<td>61.41 (7.84)</td>
<td>0.000</td>
<td>0 (0-5.6)</td>
<td>2.30 (2.14)</td>
<td>p = 0.20</td>
<td>0.82 (0.06)</td>
<td>***</td>
<td>0.10 (0.06)</td>
<td>** 0.37 (0.87)</td>
<td></td>
</tr>
<tr>
<td>1yr</td>
<td>M1</td>
<td>2677</td>
<td>2689</td>
<td>0 (0)</td>
<td>74.29 (8.62)</td>
<td>0.000</td>
<td>0 (0-3.3)</td>
<td>16.84 (0.45)</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M2</td>
<td>2616</td>
<td>2631</td>
<td>0 (0)</td>
<td>64.57 (8.03)</td>
<td>0.000</td>
<td>0 (0-3.1)</td>
<td>5.1 (1.62)</td>
<td>** 0.49 (0.07)</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M3</td>
<td>2584</td>
<td>2607</td>
<td>0 (0)</td>
<td>61.87 (7.86)</td>
<td>0.000</td>
<td>0 (0-2.8)</td>
<td>2.0 (2.18)</td>
<td>p = 0.93</td>
<td>0.48 (0.06)</td>
<td>***</td>
<td>0.14 (0.04)</td>
<td>*** 0.86 (0.87)</td>
<td>p = 0.33</td>
</tr>
<tr>
<td>Disability</td>
<td>Post</td>
<td>M1</td>
<td>2764</td>
<td>2776</td>
<td>2.2E-10 (1.5E-05)</td>
<td>93.76 (9.68)</td>
<td>2.3E-12</td>
<td>0.000 (0-5.8)</td>
<td>20.34 (0.50)</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M2</td>
<td>2541</td>
<td>2557</td>
<td>0.90 (0.95)</td>
<td>50.24 (7.08)</td>
<td>0.018</td>
<td>1.79</td>
<td>0.78 (1.20)</td>
<td>p = 0.51</td>
<td>0.77 (0.04)</td>
<td>***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M3</td>
<td>2514</td>
<td>2537</td>
<td>0.85 (0.92)</td>
<td>49.43 (7.03)</td>
<td>0.017</td>
<td>1.71 (0-9.7)</td>
<td>2.45 (1.75)</td>
<td>p = 0.16</td>
<td>0.76 (0.04)</td>
<td>***</td>
<td>0.10 (0.03)</td>
<td>** 1.03 (0.78)</td>
</tr>
<tr>
<td>6mts</td>
<td>M1</td>
<td>2816</td>
<td>2828</td>
<td>0.80 (0.90)</td>
<td>109.29 (10.45)</td>
<td>0.007</td>
<td>0.73 (0-6.1)</td>
<td>18.98 (0.62)</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M2</td>
<td>2631</td>
<td>2646</td>
<td>0.93 (0.96)</td>
<td>65.25 (8.08)</td>
<td>0.014</td>
<td>1.42</td>
<td>−0.78 (1.36)</td>
<td>0.78 (0.05)</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M3</td>
<td>2574</td>
<td>2598</td>
<td>1.4E-08 (0.0001)</td>
<td>61.74 (7.86)</td>
<td>2.3E-10</td>
<td>0.000</td>
<td>−7.61 (1.89)</td>
<td>0.76 (0.05)</td>
<td>***</td>
<td>0.20 (0.04)</td>
<td>** 1.48 (0.88)</td>
<td>p = 0.09</td>
<td></td>
</tr>
<tr>
<td>1yr</td>
<td>M1</td>
<td>2878</td>
<td>2889</td>
<td>2.1E-07 (0.00005)</td>
<td>127.08 (11.27)</td>
<td>1.7E-10</td>
<td>0.000 (0-7.1)</td>
<td>17.9 (0.58)</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M2</td>
<td>2750</td>
<td>2766</td>
<td>1.25 (1.12)</td>
<td>88.19 (9.39)</td>
<td>0.014</td>
<td>1.42 (0-8.5)</td>
<td>−0.42 (1.58)</td>
<td>p = 0.57</td>
<td>0.72 (0.05)</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M3</td>
<td>2697</td>
<td>2720</td>
<td>1.25 (1.12)</td>
<td>81.63 (9.03)</td>
<td>0.015</td>
<td>1.53 (0-9.4)</td>
<td>−8.81 (2.25)</td>
<td>***</td>
<td>0.70 (0.06)</td>
<td>*** 0.24 (0.04)</td>
<td>*** −1.04 (1.01)</td>
<td>p = 0.31</td>
<td></td>
</tr>
</tbody>
</table>

M1 = unconditional model; M2 = model conditional to baseline scores; M3 = conditional model to baseline score with covariates (age and gender). AIC = Akaike’s Information Criterion; BIC = Bayesian information criteria; 95% CI = 95% credible interval. **p < .001; *p < .05.
outpatient clinic by similarly qualified staff. Although almost never accounted for by research designs, due to measurement complexity, the environment and the “non-therapeutic services” associated with attending a therapy session (e.g. admin staff, reception area) may have an effect on therapeutic outcomes. Multi-centre studies may therefore suffer from the effect of unshared environment compared to single centre studies.

In our study the therapists received a comparable amount of supervision from the same supervisors, and training and professional development were also very similar. However, the number of years of experience varied greatly among the unit’s therapists (i.e. from 2 to 17 years). This suggests that appropriate training and supervision, within certain psychotherapeutic settings, could be more important than experience. Some recent research findings indicate, in line with the above speculation, that ad-hoc trained nurses, with no prior experience in psychotherapy, can deliver an effective brief CBT-based intervention for irritable bowel syndrome (Kennedy et al., 2006).

It is also interesting to note that a small proportion of therapist variance emerged in the modelling conducted with disability, suggesting that individual differences may have a different impact on different outcomes. An alternative explanation for this slight discrepancy between the fatigue and disability findings in relation to therapist effects may be the item rating different modalities of the two scales. The fatigue scores allowed four rating points while the WSAS allowed as many as eight. This difference may have produced more variability in the patients’ ratings.

The current results are thus not in line with the majority of studies from other patient groups that have reported significant therapist effects (e.g. Huppert et al., 2001; Kim et al., 2006; Luborsky et al., 1997). One exception is the study by Elkin et al. (2006) that did not find support for any significant therapist effects in their analysis of National Institute of Mental Health Treatment of Depression Collaborative Research Program. Our results support Elkin et al. (2006), and also strengthen the validity of Crits-Cristoph and Gallop’s assertion that therapist effects may be overestimated in the literature (2006).

The sample size of the therapists mainly determines the power and the accuracy of intraclass correlation estimates. The relatively small number of therapists in this study, therefore, limits the accuracy of our estimates of therapist variability, which resulted in 95% credible intervals ranging from 0 up to almost 10%. However, the results were consistent in that there were no significant differences in any of the patient outcomes between the therapists. Possible explanations as to why no therapist effects were found could involve the outcome measures used. Similar studies from other patient populations found that therapist effects vary greatly depending on the outcome measure considered (Huppert et al., 2001; Kim et al., 2006).

Therapist differences (e.g. clinical experience) might therefore be more important for other variables not included in this study (e.g. psychological distress, life satisfaction, etc.). However, the outcome measures included are widely used and accepted primary outcomes in the treatment of CFS (Price, Mitchell, Tidy, & Hunot, 2008; White, Sharpe, Chalder, DeCesare, & Walwyn, 2007).

The results, if replicated, could therefore have implications for management and treatment of CFS patients. If long clinical training and experience are not crucial for producing significant clinical changes in fatigue symptoms and disability associated with CFS, new possibilities for CBT therapists emerge. For instance, other healthcare workers such as general practitioners, nurses, occupational therapists and physiotherapists could be trained in relatively short training periods. In addition newly trained therapists, without lengthy clinical experience, could successfully treat CFS patients with CBT.

Future studies with a higher number of therapists and patients and a broader range of outcome measures are needed to replicate these results.

Acknowledgements

We would like to thank the patients and the therapists who took part in this study, in particular S.R., B.B. and A.D. We would also like to thank two anonymous reviewers for their helpful comments on this manuscript’s initial submission. Trudie Chalder acknowledges financial support from the Department of Health via the National Institute for Health Research (NIHR) Specialist Biomedical Research Centre for Mental Health award to South London and Maudsley NHS Foundation Trust (SLaM) and the Institute of Psychiatry at King’s College London.

References


