
Article:
Delgadillo, J. orcid.org/0000-0001-5349-230X, Saxon, D. and Barkham, M. (2018) Associations between therapists' occupational burnout and their patients' depression and anxiety treatment outcomes. Depression and Anxiety, 35 (9). pp. 844-850. ISSN 1091-4269

https://doi.org/10.1002/da.22766

Associations between therapists' occupational burnout and their patients' depression and anxiety treatment outcomes

Jaime Delgadillo Ph.D. [1,2], David Saxon Ph.D.[2], and Michael Barkham Ph.D.[1,2]

1. Clinical Psychology Unit, Department of Psychology, University of Sheffield, Sheffield, United Kingdom
2. Centre for Psychological Services Research, School of Health and Related Research, University of Sheffield, Sheffield, United Kingdom

Declarations of interest: None.

Correspondence: jaime.delgadillo@nhs.net

Abstract

Background: Occupational burnout is common in mental health professionals, but its impact on patient outcomes is as yet uncertain. This study aimed to investigate associations between therapist-level burnout and patient-level treatment outcomes after psychological therapy.

Methods: We applied multilevel modelling using depression (PHQ-9) and anxiety (GAD-7) outcomes data from 2223 patients nested within 49 therapists. Therapists completed a survey including the Oldenburg Burnout Inventory (OLBI) and a job satisfaction scale (JDSS).

Results: After controlling for case-mix, around 5% of variability in treatment outcomes was explained by *therapist effects* (TE). Higher therapist OLBI-Disengagement and JDSS scores were significantly associated with poorer treatment outcomes, explaining between 31% and 39% of the TE estimate. Higher OLBI scores were also correlated with lower job satisfaction ratings.

Conclusions: Therapist burnout has a negative impact on treatment outcomes and could be the target of future preventive and remedial action.

Key words: burnout; therapist effects; multilevel modelling; psychological therapies

1. INTRODUCTION

Like many other roles that involve helping people, providing psychological therapy can be a challenging and emotionally taxing line of work. For example, vicarious traumatization and secondary traumatic stress in mental health professionals are well-documented adverse reactions arising from work with trauma survivors (Baum, 2016; Canfield 2005; Sabin-Farrell & Turpin, 2003). Secondary traumatic stress symptoms are similar to those of post-traumatic stress disorder (PTSD) (Figley, 1995), whereas vicarious trauma primarily involves cognitive changes related to processing disturbing accounts of clients' traumatic experiences (Sabin-Farrell & Turpin, 2003). More broadly, and beyond the specific domain of PTSD treatment, compassion fatigue denotes a state of emotional exhaustion that can occur as a result of intensive empathic involvement with people who are in distress (Figley, 2002). Compassion fatigue has been identified as a commonly occurring reaction in mental health professionals (Ray, Wong, White, & Heaslip, 2013; Rossi et al., 2012). Secondary traumatic stress, vicarious traumatization and compassion fatigue have all been recognized as pathways to occupational burnout (Canfield 2005; Ray et al., 2013; Rossi et al., 2012; Sabin-Farrell & Turpin, 2003). Burnout in mental healthcare has also been linked to wider organisational factors such as increased workload, time pressures, safety issues, role ambiguity, lack of supervision and reduced resources (Edwards, Burnard, Coyle, Fothergill, & Hannigan, 2000). Altogether, these studies demonstrate the numerous emotional difficulties that may be experienced by mental health professionals. Maslach (1982) referred to these adverse reactions as *the cost of caring*.

It has been estimated that between 21% and 67% of mental health workers experience occupational burnout (Morse, Salyers, Rollins, Monroe-DeVita, & Pfahler, 2012), with adverse consequences for their general health (Acker, 2010),

attitudes towards clients (Holmqvist, & Jeanneau, 2006), and job satisfaction (Blankertz, & Robinson, 1997). In particular, therapists working in institutional settings appear to be at increased risk of burnout compared to those in private practice (Farber, 1985). Recent surveys of clinicians working in publicly funded psychological therapy services revealed that higher burnout was associated with increased job demands, stymied autonomy, greater in-session anxiety, increased hours of over-time work, high volume of telephone-based work and fewer hours of clinical supervision (Steel, Macdonald, Schröder, & Mellor-Clark, 2015; Westwood, Morison, Allt, & Holmes, 2017). In these circumstances, it is plausible that therapist burnout could be associated with poorer treatment outcomes, although there is as yet no published evidence in support of this hypothesis. To address this gap in the literature, the current study investigated potential associations between therapist-level burnout and patient-level measures of treatment outcomes in a primary care psychological therapy setting.

2. MATERIALS AND METHODS

2.1. Setting, interventions and study design

This study was conducted in a publicly funded psychological therapy service in northern England, which was part of the national *Improving Access to Psychological Therapies* (IAPT) programme (Clark, 2011). IAPT services offer evidence-based psychological interventions for depression and anxiety problems organised in a stepped care model. These include low intensity (<8 sessions) guided self-help based on principles of cognitive behavioural therapy (CBT), and high intensity (up to 20 sessions) psychotherapeutic interventions including CBT, interpersonal psychotherapy, counselling for depression, and eye-movement desensitization and reprocessing (EMDR) for PTSD. Most patients initially access low intensity

interventions, and high intensity interventions are offered to patients with more severe symptoms or those who do not benefit from the initial step.

Three groups of clinicians were included in the study: (1) psychological wellbeing practitioners (PWP) trained to a postgraduate certificate in low intensity CBT; (2) cognitive behavioural therapists trained to a postgraduate diploma in high intensity CBT; and (3) mental health nurses (MHN) with post-qualification training to deliver other high intensity treatments listed above. Participating therapists provided informed consent to complete an electronic survey and to link the responses to clinical information from their patients in a fully anonymized dataset. The study was approved by an NHS research ethics committee (REC Ref: 11/YH/0005).

The primary objective of the study was to determine if therapist-level burnout was significantly associated with patient-level outcomes, after controlling for case-mix. A secondary objective was to explore potential predictors of occupational burnout.

2.2. Measures and data sources

Therapist-level data

An electronic survey was used to gather de-identified data on therapists' age, gender, ethnicity, years of experience delivering psychological care, role in the service, along with validated measures of occupational burnout and job satisfaction.

Occupational burnout was measured using the Oldenburg Burnout Inventory (OLBI), a 16-item questionnaire designed to assess two facets of burnout, emotional *exhaustion* (OLBI-E) and *disengagement* (OLBI-D), including their cognitive and somatic aspects (Demerouti, Bakker, Nachreiner, & Schaufeli, 2001). For both dimensions, four items are phrased positively and four items are phrased

negatively (reverse scored). Every item is scored between 1 (strongly agree) and 4 (strongly disagree), and item ratings are averaged into a single index (range = 1 to 4), where a higher score is indicative of increased burnout. Examples of positively and negatively phrased items are: "I can tolerate the pressure of my work very well"; "During my work, I often feel emotionally drained". A psychometric validation study applying the OLBI in 2599 adults with a variety of professional backgrounds demonstrated high internal consistency (Cronbach's alpha = .74 to .76 for each subscale) as well as robust convergent, and discriminant validity (Halbesleben & Demerouti, 2005). Cronbach's alpha indices in this study sample were OLBI-D α = .87 and OLBI-E α = .84.

Job satisfaction was assessed using the Job Discrepancy and Satisfaction Scale (JDSS; Nagy, 2002). This 5-item measure addresses the extent to which practitioners are satisfied with their current working conditions including work tasks, pay, promotions, supervision and co-workers. Items are scored on a scale between 1 (not at all satisfying) and 4 (very satisfying), where a higher mean score (range = 1 to 4) is indicative of greater job satisfaction, with a minimum reliability (item correlation) estimate of r = .63. Cronbach's alpha in this study sample was α = .75.

Patient-level data

Patients accessing this service completed two standardised outcome measures on a session-to-session basis to monitor response to treatment. The PHQ-9 is a nine-item screening tool for major depression, where each item is rated on a 0 to 3 scale, yielding a total depression severity score between 0–27 (Kroenke, Spitzer, & Williams, 2001). A cut-off ≥ 10 has been recommended to detect clinically significant depression symptoms. The GAD-7 is a seven-item measure developed to screen for anxiety disorders (Kroenke, Spitzer, Williams, Monahan, & Löwe,

2007). It is also rated using a 0 to 3 scale, yielding a total anxiety severity score between 0–21. A cut-off score ≥8 is recommended to identify the likely presence of a diagnosable anxiety disorder. The validity and reliability of both measures are well-established across different countries and healthcare populations, with pooled (across multiple studies) sensitivity and specificity indices upwards of .78 (Moriarty, Gilbody, McMillan, & Manea, 2015; Plummer, Manea, Trepel, & McMillan, 2016).

Additional patient data included demographics (age, gender, ethnicity, employment status, socioeconomic status) and clinical information (diagnosis, treatments received, functional impairment, comorbidity of medical long-term conditions, use of antidepressants). Socioeconomic status was measured using the English index of multiple deprivation (IMD; Department for Communities and Local Government, 2011), split into quintile groups. Functional impairment was measured using the Work and Social Adjustment Scale (WSAS), where each of five items is rated on a scale of 0 (no impairment) to 8 (very severe impairment), rendering a total functional impairment score between 0–40 (Mundt, Marks, Shear, & Greist, 2002). Comorbid long-term conditions (i.e., diabetes, asthma, chronic pain, cardiovascular disease, etc.) were screened using a standardized checklist administered at referral (Delgadillo, Dawson, Gilbody, & Böhnke, 2016).

Therapist sample

A total of 56 therapists provided consent, participated in the electronic survey and had their responses linked to their patients' data (N = 3728). Patients who attended only one session (N = 726) or did not provide pre and post therapy outcomes measures (N = 402) where excluded. Therapists who treated less than 20 cases within one year after completing the survey were excluded (N = 7; 12.5%), in line with minimum sample size recommendations for multilevel modelling (Schiefele et

al., 2017). This resulted in a study sample of 2509 patients nested within 49 therapists.

2.3. Statistical analysis

The analysis plan was organised in two stages, consistent with the study objectives outlined above. In stage 1, we applied multilevel modelling (MLM), including patient-level case-mix variables (level 1) and therapist-level predictors (level 2). Pre-post treatment change scores (positive scores denoting improvement) in PHQ-9 and GAD-7 were taken as the dependent variables, with separate models for each outcome. MLM was applied in several steps. First, a single-level case-mix model was developed to identify statistically significant patient characteristics. Continuous variables were grand-mean centred, and significant non-linear relationships between independent variables and outcomes were modelled using polynomial (e.g., quadratic) terms. Once an optimal case-mix model was obtained, a level-2 random intercept was fitted and random slopes were also examined. Improvements in model fit were assessed by -2 log likelihood ratio tests. The intracluster correlation coefficient (ICC) was estimated as a baseline measure of variability in outcomes explained by differences between therapists (e.g., *therapist effects*). Next, therapist variables were introduced as level-2 predictors (OLBI-D, OLBI-E, JDSS, age, gender, ethnicity, years or experience, role, caseload size) after which non-significant variables were removed to obtain a parsimonious model. All categorical variables (e.g. gender, role) were coded as dummy variables. The ICC was re-calculated after including only statistically significant level-2 predictors. The difference between the first (case-mix adjusted MLM without level-2 predictors) and second (including significant level-2 predictors) ICC calculation as estimated. The reduction of the ICC statistic indicates that the fully adjusted model explains a greater proportion of variance in treatment outcomes, resulting from the

inclusion of relevant level-2predictors. Therefore, this difference in ICC gives us an indication of the proportion of the *therapist effect* that is explained by the level-2 predictors.

The primary analysis was applied in a dataset where complete data were available for all predictor variables (N = 2223). Additionally, we repeated the above MLM strategy as a sensitivity analysis in a dataset where missing values were imputed by aggregating 25 iterations using the Markov Chain Monte Carlo (MCMC) method. IMD scores could not be imputed, therefore the sensitivity analysis was carried out on a sample of 2393, 116 less than the original sample.

Therapist residuals with 95% confidence intervals were ranked to produce a caterpillar plot, denoting the degree to which each therapist's outcomes depart from that of the 'average therapist', after controlling for case-mix (Goldstein & Spiegelhalter, 1996; Saxon & Barkham, 2012). These plots were derived from multilevel models that did not adjust predicted outcomes for OLBI scores. This enabled us to visually examine relationships between therapist rankings and raw OLBI scores.

In stage 2, we used a summarised therapist-level dataset, where case-mix variables were averaged within each caseload (e.g., mean baseline PHQ-9 scores to denote average caseload severity). Ordinary least squares regression was applied to examine therapist and caseload variables as predictors of OLBI.

3. RESULTS

3.1. Sample characteristics

The characteristics of included therapists are summarised in Table 1. Comparing the three groups of therapists defined by their roles (treatment modality), we found significant differences in mean age, years of experience and caseload size derived

from all patients seen (Kruskal-Wallis tests, all *p*-values < 0.001). On average, the PWP group included younger practitioners, with fewer years of experience and considerably larger caseloads. The CBT group tended to have a higher proportion of male therapists and smaller caseloads. No significant group-level differences were observed in mean OLBI (F[48] = 3.25, p = 0.05) or JDSS measures (F[53] = 1.10, p = 0.34).

[Table 1]

Overall, the patient sample was characterised by a majority of female (65.5%) patients with a mean age of 38.40 (SD = 13.40) and from a white British background (89.4%). Approximately 12.1% had a comorbid long-term health condition and 24.2% were unemployed. The distribution of cases across IMD quintiles was: Q1 (most deprived areas) = 23.5%, Q2 = 21.3%, Q3 = 21.4%, Q4 = 20.1%, Q5 (most affluent areas) = 13.8%. The three most frequent diagnoses recorded in clinical records were depression (34.3%), mixed anxiety and depressive disorder (33.6%), and generalized anxiety disorder (11.1%). Mean baseline severity estimates were PHQ-9 = 13.87 (SD = 6.66), GAD-7 = 12.37 (SD = 5.55) and WSAS = 18.55 (SD = 9.53). More than half (53.9%) of all patients had been prescribed antidepressants.

3.2. Associations between case-mix variables and clinical outcomes

The fully-adjusted MLM equations are presented in Table 2. Due to missing patient data for some of the predictor variables, the final models included 2223 patients, with the 49 therapists having between 12 and 106 patients, and a mean (SD) number of patients per therapist of 45.4 (23.85). The models for both outcomes were broadly similar, with poorer outcomes associated with higher baseline WSAS

scores and membership of a minority ethnic group. Also, those who were unemployed and those living in more deprived neighbourhoods tended to have poorer outcomes. In both the PHQ-9 and GAD-7 models, higher baseline severity (PHQ-9 and GAD-7 respectively) tended to be associated with greater improvement, although this is largely a statistical artefact (ceiling and floor effects) since cases with higher baseline scores have greater room for improvement. However, curvilinear (quadratic) relationships were observed for baseline PHQ-9 with both outcomes, indicating that improvement in anxiety and depression symptoms was diminished in cases with the highest initial depression severity. In contrast, GAD-7 baseline score was not a predictor of PHQ-9 change. There was also a significant random slope for baseline PHQ-9 in the depression model and baseline GAD-7 in the anxiety model indicating that there was greater variability in treatment outcomes between therapists as intake severity increased.

[Table 2]

3.3. Associations between therapist variables and clinical outcomes

After adjusting for relevant case-mix variables, multilevel models that included random intercepts yielded therapist effects of 5.7% (PHQ-9 model) and 5.6% (GAD-7 model). Of the available level-2 variables, the therapist OLBI-D score was a significant predictor, where higher scores were associated with less symptomatic improvement in depression and anxiety. OLBI-E was not significant in either model. Therapist JDSS was a significant predictor of PHQ-9 but not GAD-7 outcomes. In the PHQ-9 model JDSS reduced the therapist effect from 5.7% to 5.1% while OLBI-D reduced it further to 3.5%. In the GAD-7 model OLBI-D reduced the therapist effect from 5.6% to 4.0%. Therefore these therapist variables explained approximately 38.6% and 28.6% of the therapist effects respectively.

Although the random slopes significantly improved model fit in both models, there was some uncertainty regarding the extent of the differences in slopes between therapists in the GAD-7 model, as indicated by the large standard error (coefficient: 0.005; SE: 0.003).

Sensitivity analyses using the imputed dataset tended to reduce the size of model coefficients generally, which resulted in patient ethnicity and therapist JDSS score no longer being significant in the PHQ-9 model. In addition, the random slopes were no longer significant in both models. However, the predictor variables were broadly similar and, overall, sensitivity analyses indicated that the predictive value of OLBI-D was robust to changes in sample size and missing data (see Table 3). As in the primary analysis, OLBI-D explained around 30% of the *therapist effects*.

[Table 3]

[Figure 1]

Figure 1 presents caterpillar plots where therapist residuals are ranked from least (left) to most effective (right). Regarding depression outcomes (Panel A), 10 therapists were ranked as significantly below average (i.e. their 95% CIs did not cross zero) and 6 were better than average. There was less variability in anxiety outcomes (Panel B), with 2 therapists ranked as below and 3 as above average. Therapists are identified as having low (green), medium (grey) or high (red) OLBI-D scores, defined based on scores above or below 1 standard deviation of the mean (M = 2.15, SD = 0.52; ≤1.62 = low, 1.63 to 2.67 = medium, ≥2.68 = high). A general trend is evident in both panels, where therapists with lower than average (green) OLBI-D scores tend to be located towards the right side (more effective).

3.4. Correlates of therapist disengagement

A linear regression analysis indicated that OLBI-D scores were significantly associated ($p < .05$) with the therapist's role (B = 0.174, SD = 0.061, 95% CI = 0.025 to 0.323) and JDSS scores (B = -0.533, SD = 0.140, 95% CI = -0.814 to -0.252). On average, MHN practitioners tended to have the highest OLBI-D scores (M = 2.34, SD = 0.52), followed by PWPs (M = 2.29, SD = 0.47), and CBTs (M = 1.92, SD = 0.37). The inverse correlation between OLBI-D and JDSS was in the magnitude of $r = -0.51$, $p < 0.001$. Overall, role explained approximately 13% and JDSS explained approximately 34% of variance in OLBI-D scores. No other therapist or caseload characteristics were associated with OLBI-D scores.

4. DISCUSSION

4.1. Main findings

Consistent with our hypothesis, we found that therapists' occupational burnout levels were associated with their patients' psychological treatment outcomes measured using patient-reported depression and anxiety scales. Lower job satisfaction was also associated with poorer depression outcomes, but not with anxiety outcomes. Therapists with lower indices of burnout tended to cluster among the more effective therapists using a case-mix adjusted ranking method (caterpillar plots). We found that specifically the disengagement domain of the OLBI measure was associated with treatment outcomes, but the exhaustion domain was not. This aspect of disengagement, which is theoretically akin to Maslach's notion of *depersonalisation* (Maslach, Jackson, & Leiter, 1996), is likely to be particularly important in psychological treatment. Stressful disengagement may be plausibly related to an impaired ability to express empathy and to form an

effective working alliance with patients. Both empathy and alliance are well-established predictors of treatment outcomes (Wampold, 2015), and occupational burnout may potentially mitigate improvement through the dampening of these common therapy processes. If occupational burnout is the *cost of caring* for therapists (Maslach, 1982), then poor treatment outcomes is the *cost of disengagement*.

The overall magnitude of *therapist effects* (TE) in this sample was in the region of 5%, which is comparable although marginally smaller to that observed in other studies conducted in IAPT services (Firth, Barkham, Kellett, & Saxon, 2015; Green, Barkham, Kellett, & Saxon, 2014; Pereira, Barkham, Kellett, & Saxon, 2017). Given the reduction of the magnitude of TE after controlling for burnout, it is plausible that outcome differences between therapists may be partly explained by differences in occupational stress and coping resources.

We found no evidence that caseload-specific factors (e.g., caseload size or severity) influence burnout levels, although such aggregated metrics may not adequately capture aspects of stressful work involvement identified in other studies (Steel et al., 2015). Equally, therapist role, years of experience, ethnicity, age, etc., were not associated with treatment outcomes. As expected, higher indices of burnout were correlated with lower job satisfaction ratings. Furthermore, mental health nurse practitioners tended to have higher indices of burnout compared to the other occupational groups, and our regression analyses suggest that this was unlikely to be explained by other variables such as caseload size or caseload severity. It may be that differences in the relative esteem, pay, supervisory quality or other aspects of this role may increase this group's higher propensity towards burnout. We note that the role variable was not associated with treatment outcomes when OLBI and JDSS variables were included in the MLM analysis. This suggests that a therapist's role (e.g., therapeutic orientation, occupational group)

is not associated with treatment outcomes, but certain roles may be more prone to burnout and lower job satisfaction – which are more directly associated with treatment outcomes.

A previous study in a similar setting has also indicated that wider organizational and contextual factors such as working over-time, increased telephone-based work and infrequent access to clinical supervision are implicated in occupational burnout (Westwood et al., 2017). Therapists' resilience may also be an important determinant of clinical outcomes (Green et al., 2014; Pereira et al., 2017). From a theoretical perspective, burnout and resilience could be seen as related factors, such that one may be moderated by the other, and could be plausibly influenced by individual ways of coping that may serve to either mitigate or maintain burnout (Tyrrell, 2010).

4.2. Methodological considerations

To our knowledge, this is the first *therapist effects* study to empirically test the influence of therapists' occupational burnout on observed clinical outcomes, using a large sample and appropriate analytical procedures (MLM) which take account of the nested structure of patient and therapist variables. In spite of the availability of data for over 2000 patients nested within 49 therapists, the sample size (particularly the number of patients per therapist) may not be sufficient to produce reliable estimates of all model parameters. A further limitation concerns the cross-sectional survey method, which does not enable us to make conclusive inferences about the direction of causality. It is plausible that occupational burnout and job satisfaction vary over time. Future replication of these findings is necessary, particularly in longitudinal designs applying repeated measurement of occupational burnout and job satisfaction. This would enable us to assess the

stability of these measures and also to better understand the temporal relationships between therapist burnout and patients' outcomes.

A plausible interpretation of our findings is that more stressful work conditions (reduced autonomy, working over-time, infrequent clinical supervision, coping deficits) could increase burnout, which in turn attenuates clinical improvement through the mechanism of stressful disengagement and its influence on the therapeutic alliance. An alternative explanation could be that the observation of poorer clinical outcomes (e.g., due to case complexity or competency deficits) could demoralise and lead therapists to become burned out over time. Future studies with longitudinal designs are necessary to better understand the interrelationships between case-mix, organizational context, occupational burnout, job satisfaction and clinical outcomes. Another caveat to note is that this study was carried out within the context of stepped care, in an IAPT service with a workforce that may not be representative of that in other settings (e.g., hospital or secondary care services).

5. CONCLUSION

Therapists' occupational burnout is associated with poorer psychological treatment outcomes. This and other recent studies described above raise a growing concern around therapists' wellbeing in publicly funded mental health services. Future directions to address this concern may involve both organizational re-design and interventions to enhance coping and resilience for mental healthcare practitioners, and particularly to support those with a propensity towards occupational burnout.

References

Acker, G. M. (2010). The challenges in providing services to clients with mental illness: Managed care, burnout and somatic symptoms among social workers. *Community Mental Health Journal, 46*, 591-600.

Baum, N. (2016). Secondary traumatization in mental health professionals: a systematic review of gender findings. *Trauma, Violence, & Abuse*, *17,* 221-235.

Blankertz, L. E., & Robinson, S. E. (1997). Turnover intentions of community mental health workers in psychosocial rehabilitation services. *Community Mental Health Journal, 33*, 517-529.

Canfield, J. (2005). Secondary traumatization, burnout, and vicarious traumatization: A review of the literature as it relates to therapists who treat trauma. *Smith College Studies in Social Work*, *75*, 81-101.

Clark, D. M. (2011). Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: the IAPT experience. *International Review of Psychiatry, 23*, 318-327.

Delgadillo, J., Dawson, A., Gilbody, S., & Böhnke, J. R. (2016). Impact of long-term medical conditions on the outcomes of psychological therapy for depression and anxiety. *The British Journal of Psychiatry*, doi: 10.1192/bjp.bp.116.189027.

Demerouti, E., Bakker, A.B., Nachreiner, F., Schaufeli, W.B. (2001). The job demands-resources model of burnout. *Journal of Applied Psychology, 86*, 499–512.

Department for Communities and Local Government. (2011). *The English indices of deprivation 2010*. London: Author.

Edwards, D., Burnard, P., Coyle, D., Fothergill, A., & Hannigan, B. (2000). Stress and burnout in community mental health nursing: a review of the literature. *Journal of Psychiatric and Mental Health Nursing, 7*, 7-14.

Farber, B.A. (1985). Clinical psychologists' perceptions of psychotherapeutic work. *Clinical Psychologist, 38*, 10-13.

Figley, C.R. (1995). Compassion fatigue as secondary traumatic stress disorder: An overview. In C. R. Figley (Ed.), *Compassion fatigue: Coping with secondary traumatic stress disorder in those who treat the traumatized*. New York, NY: Brunner/Mazel, pp. 1-20.

Figley, C.R. (2002). Compassion fatigue: Psychotherapists' chronic lack of self care. *Journal of Clinical Psychology*, *58*,1433-41.

Firth, N., Barkham, M., Kellett, S., & Saxon, D. (2015). Therapist effects and moderators of effectiveness and efficiency in psychological wellbeing practitioners: a multilevel modelling analysis. *Behaviour Research and Therapy, 69*, 54-62.

Goldstein, H., & Spiegelhalter, D. J. (1996). League tables and their limitations: statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society - Series A*, *159*, 385-443.

Green, H., Barkham, M., Kellett, S., & Saxon, D. (2014). Therapist effects and IAPT Psychological Wellbeing Practitioners (PWPs): A multilevel modelling and mixed methods analysis. *Behaviour Research and Therapy, 63*, 43-54.

Halbesleben, J. R., & Demerouti, E. (2005). The construct validity of an alternative measure of burnout: Investigating the English translation of the Oldenburg Burnout Inventory. *Work & Stress, 19*, 208-220.

Holmqvist, R., & Jeanneau, M. (2006). Burnout and psychiatric staff's feelings towards patients. *Psychiatry Research, 145*, 207-213.

Kroenke, K., Spitzer, R.L., & Williams, J.B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*, 606-613.

Kroenke, K., Spitzer, R.L., Williams, J.B.W., Monahan, P.O., & Löwe, B. (2007). Anxiety disorders in primary care: Prevalence, impairment, comorbidity, and detection. *Annals of Internal Medicine*, *146*, 317–325.

Maslach, C. (1982). *Burnout: The cost of caring*. Los Altos, CA: ISHK.

Maslach, C., Jackson, S.E., & Leiter, M.P. (1996). *Maslach Burnout Inventory Manual* (3rd ed.). Palo Alto, CA.

Moriarty, A. S., Gilbody, S., McMillan, D., & Manea, L. (2015). Screening and case finding for major depressive disorder using the Patient Health Questionnaire (PHQ-9): a meta-analysis. *General Hospital Psychiatry, 37*, 567-576.

Morse, G., Salyers, M. P., Rollins, A. L., Monroe-DeVita, M., & Pfahler, C. (2012). Burnout in mental health services: A review of the problem and its remediation. *Administration and Policy in Mental Health and Mental Health Services Research, 39*, 341-352.

Mundt, J. C., Marks, I. M., Shear, M. K., & Greist, J. M. (2002). The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *The British Journal of Psychiatry, 180*, 461-464.

Nagy, M. S. (2002). Using a single-item approach to measure facet job satisfaction. *Journal of Occupational and Organizational Psychology, 75*, 77-86.

Pereira, J. A., Barkham, M., Kellett, S., & Saxon, D. (2017). The role of practitioner resilience and mindfulness in effective practice: A practice-based feasibility study. *Administration and Policy in Mental Health and Mental Health Services Research, 44*, 691-704.

Plummer, F., Manea, L., Trepel, D., & McMillan, D. (2016). Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic metaanalysis. *General Hospital Psychiatry, 39*, 24-31.

Ray, S. L., Wong, C., White, D., & Heaslip, K. (2013). Compassion satisfaction, compassion fatigue, work life conditions, and burnout among frontline mental health care professionals. *Traumatology, 19*, 255-267.

Rossi, A., Cetrano, G., Pertile, R., Rabbi, L., Donisi, V., Grigoletti, L., ... Amaddeo, F. (2012). Burnout, compassion fatigue, and compassion satisfaction among staff in community-based mental health services. *Psychiatry Research, 200*, 933-938.

Sabin-Farrell, R., & Turpin, G. (2003). Vicarious traumatization: implications for the mental health of health workers? *Clinical Psychology Review*, *23*, 449-80.

Saxon, D., & Barkham, M. (2012). Patterns of therapist variability: Therapist effects and the contribution of patient severity and risk. *Journal of Consulting and Clinical Psychology, 80*, 535-546.

Schiefele, A. K., Lutz, W., Barkham, M., Rubel, J., Böhnke, J., Delgadillo, J., ... Lambert, M. J. (2017). Reliability of therapist effects in practice-based psychotherapy research: A guide for the planning of future studies. *Administration and Policy in Mental Health and Mental Health Services Research, 44*, 598-613.

Steel, C., Macdonald, J., Schröder, T., & Mellor-Clark, J. (2015). Exhausted but not cynical: burnout in therapists working within Improving Access to Psychological Therapy Services. *Journal of Mental Health, 24*, 33-37.

Tyrrell, Z. (2010). A cognitive behavioural model for maintaining processes in burnout. *The Cognitive Behaviour Therapist, 3*, 18-26.

Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry*, *14*, 270-277.

Westwood, S., Morison, L., Allt, J., & Holmes, N. (2017). Predictors of emotional exhaustion, disengagement and burnout among improving access to psychological therapies (IAPT) practitioners. *Journal of Mental Health, 26*, 172-179.

Table 1. Therapist-level characteristics

	Full sample (N = 49)	**PWP** (N = 13)	**MHN** (N = 15)	**CBT** (N = 21)
Demographics				
Age – mean (SD)*	37.9 (10.11)	30.9 (10.16)	44.8 (9.04)	37.7 (7.61)
Females N(%)	43 (87.8)	13 (100)	14 (93.3)	16 (76.2)
White British N(%)	46 (93.9)	12 (92.3)	14 (93.3)	20 (95.2)
Occupational variables				
Years of experience – mean (SD)*	11.8 (7.02)	5.6 (4.68)	16.7 (4.79)	12.2 (6.77)
Caseload size – mean (SD)*	72.7 (35.72)	103.3 (38.58)	80.9 (29.84)	47.8 (15.33)
OLBI-D mean (SD)	2.2 (0.52)	2.3 (0.47)	2.3 (0.65)	1.9 (0.37)
OLBI-E mean (SD)	2.4 (0.51)	2.4 (0.53)	2.6 (0.53)	2.3 (0.48)
JDSS – mean (SD)	2.7 (0.45)	2.8 (0.47)	2.5 (0.49)	2.7 (0.42)

PWP = psychological wellbeing practitioners; MHN = mental health nurse practitioners; CBT = cognitive behavioural therapists; OLBI = Oldenburg burnout inventory, D = disengagement, E = exhaustion; JDSS = Job discrepancy and satisfaction scale; * significant between-group differences; Caseload size is based on all patients seen by practitioners.

Table 2. Fully adjusted multilevel models with OLBI-Disengagement scores predicting change in depression (PHQ-9) and anxiety (GAD-7)

Depression model	Anxiety model
$PHQ9\_change_{ij} = \beta_{0j} + \beta_{1j}$ (PHQ9_baseline_gmc] $_{ij}$	$GAD7\_change_{ij} = \beta_{0j} + \beta_{1j}$ (GAD7_baseline_gmc) $_{ij}$
–0.005(0.002)(PHQ9_baseline_gmc)$^2_{ij}$	–0.172(0.026)(PHQ9_baseline_gmc) $_{ij}$
–0.094(0.017)(WSAS_baseline_gmc) $_{ij}$	–0.006(0.002)(PHQ9_baseline_gmc)^2 $_{ij}$
+ 0.618(0.342)IMD_quintile2$_{ij}$	–0.056(0.015)(WSAS_baseline_gmc)$_{ij}$
+ 0.837(0.331)IMD_quintile3$_{ij}$	+ 0.463(0.300)IMD_quintile2$_{ij}$
+ 1.180(0.358)IMD_quintile4$_{ij}$	+ 0.376(0.298)IMD_quintile3$_{ij}$
+ 1.269(0.390)IMD_quintile5$_{ij}$	+ 0.957(0.321)IMD_quintile4$_{ij}$
–1.921(0.297)Unemployed$_{ij}$	+ 0.484(0.341)IMD_quintile5$_{ij}$
–0.922(0.381)Minority_ethnic_group$_{ij}$	–1.586(0.262)Unemployed$_{ij}$
-0.780(0.216)OLBI_D_gmc$_j$	–1.286(0.334)Minority_ethnic_group$_{ij}$
+0.607(0.253)JDSS_gmc$_j$	–0.881(0.160)OLBI_D_gmc$_j$
+ e_{ij}	+ e_{ij}
$\beta_{0j} = 5.282(0.318) + u_{0j}$	$\beta_{0j} = 5.155(0.286) + u_{0j}$
$\beta_{1j} = 0.516(0.027) + u_{1j}$	$\beta_{1j} = 0.676(0.028) + u_{1j}$
u_{0j} ~ N(0, Ω_u) : Ω_u = 0.997(0.307)	u_{0j} ~ N(0, Ω_u) : Ω_u = 0.895(0.262)
u_{1j} 0.133(0.035) 0.010(0.004)	u_{1j} 0.105(0.028) 0.005(0.003)
e_{ij} ~ N(0, σ^2_e) σ^2_e = 27.777(0.842)	e_{ij} ~ N(0, σ^2_e) σ^2_e = 21.634(0.656)
*–2*log-likelihood* = 13727.016	*–2*log-likelihood* = 13158.521
ICC = 0.035	*ICC* = 0.040

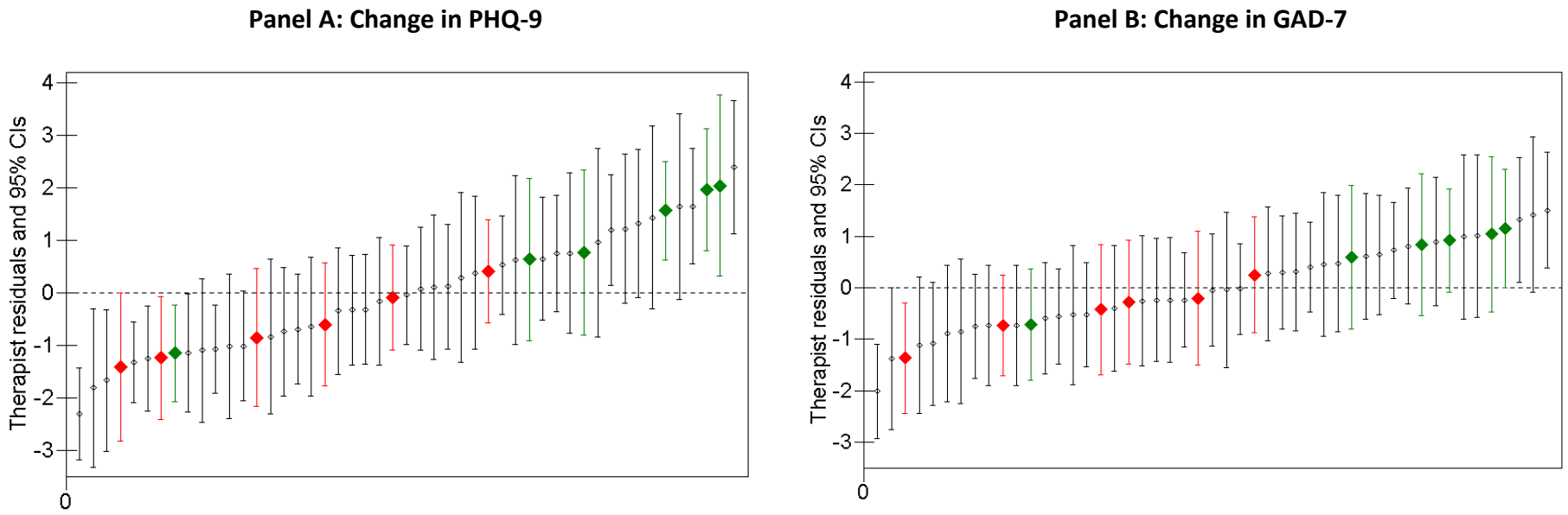
NB: PHQ-9 = depression measure; GAD-7 = anxiety measure; change = pre – post treatment change, where a positive score denotes improvement; gmc = grand-mean-centred continuous variable; WSAS = work and social adjustment scale; IMD = index of multiple deprivation, expressed as quintile groups, with most deprived quintile (1) as reference; OLBI_D = therapist-level occupational burnout measure of disengagement; ICC = intracluster correlation coefficient, as a measure of *therapist effects*

Table 3. Fully adjusted multilevel models predicting change in depression (PHQ-9) and anxiety (GAD-7) (Imputed data)

Depression model	Anxiety model
$PHQ9\_change_{ij} = \beta_{0j}$ +0.382(0.019)(PHQ9_baseline_gmc]$_{ij}$	$GAD7\_change_{ij} = \beta_{0j}$ +0.533(0.022)(GAD7_baseline_gmc)$_{ij}$
–0.007(0.002)(PHQ9_baseline_gmc)$^2_{ij}$	–0.163(0.021)(PHQ9_baseline_gmc) $_{ij}$
–0.054(0.013)(WSAS_baseline_gmc) $_{ij}$	–0.007(0.002)(PHQ9_baseline_gmc)^2 $_{ij}$
+ 0.491(0.274)IMD_quintile2$_{ij}$	–0.027(0.012)(WSAS_baseline_gmc)$_{ij}$
+ 0.762(0.277)IMD_quintile3$_{ij}$	+ 0.423(0.241)IMD_quintile2$_{ij}$
+ 1.073(0.283)IMD_quintile4$_{ij}$	+ 0.296(0.244)IMD_quintile3$_{ij}$
+ 1.260(0.326)IMD_quintile5$_{ij}$	+ 0.828(0.249)IMD_quintile4$_{ij}$
–1.816(0.225)Unemployed$_{ij}$	+ 0.574(0.286)IMD_quintile5$_{ij}$
–0.548(0.299)Minority_ethnic_group$_{ij}$	–1.502(0.198)Unemployed$_{ij}$
–1.244(0.328)OLBI_D_gmc$_j$	–0.832(0.263)Minority_ethnic_group$_{ij}$
+ e_{ij}	–0.803(0.264)OLBI_D_gmc$_j$
	+ e_{ij}
β_{0j} = 4.389(0.266) + u_{0j}	β_{0j} = 4.257(0.228) + u_{0j}
u_{0j} ~ N(0, σ^2_{u0}) σ^2_{u0} = 0.960(0.279]	u_{0j} ~ N(0, σ^2_{u0}) σ^2_{u0} = 0.578(0.183)
e_{ij} ~ N(0, σ^2_e) σ^2_e = 28.509(0.676]	e_{ij} ~ N(0, σ^2_e) σ^2_e = 22.107(0.524)
*–2*log-likelihood* = 22412.003	*–2*log-likelihood* = 21484.689
ICC = 0.033	*ICC* = 0.026

NB: PHQ-9 = depression measure; GAD-7 = anxiety measure; change = pre – post treatment change, where a positive score denotes improvement; WSAS = work and social adjustment scale; IMD = index of multiple deprivation, expressed as quintile groups, with most deprived quintile (1) as reference; Employment: category 1 = students, category 2 = unemployed, reference = employed; OLBI_D = therapist-level occupational burnout measure of Disengagement; ICC = intracluster correlation coefficient, as a measure of *therapist effects*

Figure 1. Caterpillar plots: ranking of therapists according to effectiveness and OLBI-Disengagement



NB: Therapists ranked from least (left) to most (right) effective; confidence intervals that do not overlap with dashed line are indicative of outlier therapists (least and most effective); red = therapists with OLBI-D scores that are 1 standard deviation above the mean (high burnout); green = therapists with OLBI-D scores that are 1 standard deviation below the mean (low burnout).