



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/102827/>

Version: Accepted Version

Article:

Saxon, D., Barkham, M., Foster, A. et al. (2017) The Contribution of Therapist Effects to Patient Dropout and Deterioration in the Psychological Therapies. *Clinical Psychology & Psychotherapy*, 24 (3). pp. 575-588. ISSN: 1063-3995

<https://doi.org/10.1002/cpp.2028>

This is the peer reviewed version of the following article: Saxon, D., Barkham, M., Foster, A., and Parry, G. (2016) The Contribution of Therapist Effects to Patient Dropout and Deterioration in the Psychological Therapies. *Clin. Psychol. Psychother.*, which has been published in final form at 10.1002/cpp.2028. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving (<http://olabout.wiley.com/WileyCDA/Section/id-828039.html>)

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



The contribution of therapist effects to patient dropout and deterioration in the psychological therapies

Journal:	<i>Clinical Psychology & Psychotherapy</i>
Manuscript ID	CPP-1124.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Saxon, David; University of Sheffield, ScHARR Barkham, Michael; University of Sheffield, Psychology Foster, Alexis; University of Sheffield, ScHARR Parry, Glenys; University of Sheffield, ScHARR
Keywords:	therapist effects, outcomes, drop-out, deterioration, CORE

SCHOLARONE™
Manuscripts

Review

The contribution of therapist effects to patient dropout and deterioration
in the psychological therapies

David Saxon (Centre for Psychological Services Research, University of Sheffield)*

Michael Barkham (Centre for Psychological Services Research, University of Sheffield)

Alexis Foster (School of Health & Related Research, University of Sheffield)

Glenys Parry (Centre for Psychological Services Research, University of Sheffield)

*Correspondence concerning this article should be addressed to David Saxon, School for
Health & Related Research, University of Sheffield, 30 Regents Court, Sheffield S1 4DA,
United Kingdom.

E-mail: d.saxon@sheffield.ac.uk

ABSTRACT

Background: In the psychological therapies, patient outcomes are not always positive. Some patients leave therapy prematurely (dropout) while others experience deterioration in their psychological wellbeing.

Methods: The sample for dropout comprised patients (N = 10,521) seen by 85 therapists and who attended at least the initial session of 1-to-1 therapy and completed a Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM) at pre-treatment. The sub-sample for patient deterioration comprised patients (N = 6,405) seen by the same 85 therapists but who attended 2 or more sessions, completed therapy, and returned a CORE-OM at pre- and post-treatment. Multilevel modeling was used to estimate the extent of therapist effects for both outcomes after controlling for patient characteristics.

Results: Therapist effects accounted for 12.6% of dropout variance and 10.1% of deterioration variance. Dropout rates for therapists ranged from 1.2% - 73.2%, while rates of deterioration ranged from 0% - 15.4%. There was no significant correlation between therapist dropout rate and deterioration rate (Spearman's $\rho = 0.07$, $p=0.52$).

Conclusions: The methods provide a reliable means for identifying therapists who return consistently poorer rates of patient dropout and deterioration compared to their peers. The variability between therapists and the identification of patient risk factors as significant predictors have implications for the delivery of safe psychological therapy services.

Key practitioner message:

- Therapists play an important role in contributing to patient dropout and deterioration, irrespective of case mix.
- Therapist effects on patient dropout and deterioration appear to act independently.
- Being unemployed as a patient was the strongest predictor of both dropout and deterioration
- Patient risk to self or others was also an important predictor

Keywords: Deterioration, dropout, outcomes, variability, therapist effects, CORE-OM

Introduction

Background

Clinical practice and research have, understandably, focused on the improvement patients experience when engaging in a course of psychological therapy (Lambert, 2013). However, outcomes for patients are not always positive and patients may leave therapy prematurely without making meaningful improvement (Cahill et al., 2003). Moreover, others may experience deterioration in their psychological wellbeing during the course of therapy (Craze et al., 2014; Lambert, 2010). There has to date, however, been limited research into negative outcomes in routine services and few have considered therapist effects on those outcomes. In part this has been due to the absence of sufficiently large datasets to study therapist effects, but also to inconsistencies in the definitions of the range of negative outcomes. Linden (2013) classified negative outcomes, as different types of 'unwanted events', some of which are adverse reactions to the therapy, while others may or may not be therapy related. Two manifestations of the latter are *unplanned endings*, often termed *dropout*, and *patient deterioration*.

Patient dropout

Patient dropout from therapy has been of concern in the psychological therapies for over 50 years and continues to have implications for service delivery and patient outcomes (Barrett, Chua, Crits-Christoph, Gibbons, & Thompson, 2008; Garfield, 1994; Rogers, 1951). Dropout occurs where a patient unilaterally ends therapy by ceasing to attend sessions, prior to the endpoint planned with their therapist (Westmacott, Hunsley, Best, Rumstein-McKean, & Schindler, 2010). The reported rates of dropout have ranged between 20-60% depending on the patient population, service setting, how dropout has been defined, and the methodology adopted (for details, see Reneses, Munoz, & Lopez-Ibor, 2009). A meta-

1
2
3 analysis of 669 studies of psychological and psychosocial interventions reported a dropout
4
5 rate of 17% for efficacy studies and 26% for effectiveness studies (Swift & Greenberg, 2012).
6

7
8 In the UK, successive national audits by the Royal College of Psychiatrists (RCP) of
9
10 psychological therapy services have reported treatment dropout rates of 25% and 24%
11
12 respectively (RCP, 2011, 2013), while a report on 32 UK services comprising the initial
13
14 national rollout of the Improving Access to Psychological Therapies (IAPT) initiative yielded a
15
16 rate of 21.6% (Glover, Webb, & Evison, 2010). However, these UK reports did not include
17
18 patients who failed to engage with therapy. These patients attended only one
19
20 appointment, which has been consistently found to be the modal number of psychotherapy
21
22 sessions attended (e.g., Gibbons et al., 2010). The current study considers patient dropout
23
24 at any point after the initial session.
25
26
27

28 29 *Patient deterioration*

30
31 Patient deterioration, a shorthand term for deterioration in a patient's mental state after
32
33 therapy, may be defined as any negative change between pre- and post-therapy outcome
34
35 score. Because this definition would include small changes that may be due to the inherent
36
37 unreliability of outcome measures (Jacobson & Truax, 1991), a more stringent criterion of
38
39 'statistically reliable deterioration' has been adopted by researchers (as a mirror opposite of
40
41 reliable improvement) in which measurement error is taken into account. Using this
42
43 procedure to determine rates of reliable deterioration based on selected completer samples
44
45 has yielded an estimate for primary care of 1.5% (Cahill, Barkham, & Stiles, 2010) and
46
47 upwards of 6% for secondary care (Barkham et al., 2001). Reports from the US have tended
48
49 to yield higher rates; for example, an average figure of 8.2% across a range of different
50
51 clinical settings (Hansen, Lambert, & Forman, 2002).
52
53
54
55
56
57
58
59
60

1
2
3 However, it is debatable whether the criterion for deterioration should be the same as
4
5 for improvement. The natural propensity for patient recovery, the normative trajectory of
6
7 patient change, and any statistical regression to the mean, make therapy more likely to lead
8
9 to some level of improvement rather than deterioration. In the same way that Linden
10
11 (2013) argues that if therapy does not produce the expected outcome (i.e., improvement),
12
13 then the outcome is an 'unwanted event', then reliable deterioration should not be viewed
14
15 as a mirror opposite of reliable improvement. Practitioners are likely to want to be flagged
16
17 about possible deterioration in their patients at a less stringent threshold than
18
19 improvement. Furthermore, services should be concerned if some of their practitioners
20
21 have significantly more patients who deteriorate compared to their peers, when a less
22
23 stringent threshold is used.
24
25
26
27

28 *Therapist effects*

29
30 The study of therapist effects focuses on the extent of *variability* between therapists and
31
32 the impact the individual therapists have on patient outcomes. The recommended methods
33
34 for estimating such effects, for example multilevel modeling (Goldstein & Spiegelhalter,
35
36 1996; Snijders & Bosker, 2012), require large samples of patients, and in particular
37
38 therapists (Maas & Hox, 2005). Randomised controlled trials (RCTs) are usually
39
40 underpowered to estimate therapist effects and very large datasets drawn from routine
41
42 practice are best suited to provide the statistical power and external validity needed in this
43
44 field (e.g., Castonguay, Barkham, Lutz, & McAleavy, 2013; Wampold & Brown, 2005).
45
46
47
48

49
50 Most studies of therapist effects have considered positive outcomes such as clinical
51
52 improvement or recovery rates and there is a relative paucity of research into therapist
53
54 effects on negative outcomes (Baldwin & Imel, 2013). An exception is a recent study of
55
56 patient dropout, using multilevel modeling (MLM), which found a significant therapist effect
57
58
59
60

1
2
3 (6.21%), after controlling for initial impairment, although the sample size, particularly the
4
5 number of patients per therapist, was a recognised limitation (Zimmerman, Rubel, Page &
6
7 Lutz, submitted 2016). There have been no studies to date which have used MLM to
8
9 estimate therapist effects for patient deterioration. Krause et al (2011) analysed the
10
11 outcomes for 696 therapists in the context of naturalistic treatment and found some
12
13 therapists demonstrated large, negative treatment effect sizes ($d = -0.91$ to -1.49).
14
15
16
17 However, case mix was not controlled for in the analysis.
18

19 *Case-mix*

20
21 In order to make valid comparisons between therapists' outcomes it is necessary to
22
23 control for patient characteristics that have a significant impact on outcome (i.e. case-mix).
24
25 Some likely candidates for patient dropout are: younger age (e.g., Edlund et al., 2002); non-
26
27 white ethnicity and socio-economic deprivation (e.g. Garfield, 1994) and greater intake
28
29 severity (Kazdin, Mazurick, & Siegel, 1994, Zimmerman et al, submitted 2016).
30
31
32

33
34 Few studies have considered the patient characteristics associated with deterioration
35
36 and one study failed to identify any statistically significant predictors of reliable
37
38 deterioration in a sample of 1416 UK outpatients (Shepherd, Evans, Cobb & Ghossain, 2012).
39
40 In the development of models for both dropout and deterioration, the current study will
41
42 test all available patient variables as possible case-mix variables.
43
44

45 *Study aims*

46
47 In the current study, we employed a large-scale practice-based dataset to estimate the
48
49 extent of therapist effects, while also controlling for those patient variables that have a
50
51 significant impact on outcome.
52
53

54
55 Accordingly, the study had three aims:

- 56
57 1) To estimate the therapist effect for patient dropout using MLM.
58
59
60

1
2
3 2) After applying varying indices of deterioration to the data, to estimate therapist effects
4
5 on patient deterioration for treatment completers.
6

7
8 3) To combine the variability between therapists on both dropout and deterioration and
9
10 consider whether those therapists with higher dropout rates are also those therapists with
11
12 higher deterioration rates for their treatment completers.
13
14
15
16
17
18
19

20 Method

21 *Original dataset*

22 The original data set – the Clinical Outcomes in Routine Evaluation Practice-Based Evidence
23
24 National Database-2008 – comprised information on 70,245 clients, routinely collected by
25
26 1,059 therapists in 35 UK counselling and clinical psychology services between 1999 and
27
28 2008. This data set was an updated version of earlier datasets used in studies by our
29
30 research group (e.g., Stiles, Barkham, Connell, & Mellor-Clark, 2008). Ethics approval was
31
32 covered by the UK National Health Service's Central Office for Research Ethics Committee,
33
34 application 05/Q1206/128.
35
36
37
38
39

40 *Study-specific dataset*

41 For the current study, in order to exclude practitioners who may have been selective in their
42
43 submission of patient data, therapists were only included if they provided treatment ending
44
45 information for over 90% of the patients they treated. The figure of 90% was chosen as it is
46
47 a target for the UK Improving Access to Psychological Therapies initiative (Department of
48
49 Health, 2012). Patients were included if they were 18 years old and over, were assessed and
50
51 accepted for individual therapy, completed a specified pre-therapy outcome measure (see
52
53 below), provided demographic data, and had the type of therapy ending recorded. In
54
55
56
57
58
59
60

1
2
3 addition, in order to estimate therapist variability more reliably, only therapists with 30 or
4
5 more patients were included (Soldz, 2006).
6

7
8 These criteria yielded a study-specific sample of 85 therapists and 10,521 patients from
9
10 14 sites with a range of patients per therapist of 30 – 468. In this sample, the patient mean
11
12 (SD) age was 40.3 (13.00) years, 71.2% were female, 23.9% were unemployed, and 4.6%
13
14 were of non-white ethnicity. No formal diagnoses were made but therapists recorded
15
16 patients' problems on a standardized form (CORE Assessment form; Barkham, Gilbert,
17
18 Connell, Marshall, & Twigg, 2005). This indicated that 76.8% of patients had some level of
19
20 depression (44.7% rated as ranging between moderate and severe) and 82.7% had some
21
22 level of anxiety (54.6% rated as ranging between moderate and severe).
23
24
25

26 27 *Deterioration sub-sample*

28
29 The deterioration dataset was a sub-sample of the study-specific dataset. It comprised
30
31 patients who completed therapy, had two or more sessions, and provided a pre- and post-
32
33 therapy CORE-OM score. This yielded 6,405 patients, with the same 85 therapists, who saw
34
35 between 13 –180 patients each. Therapists with less than 30 patients were not excluded, in
36
37 order to compare all 85 therapists on both outcomes. The mean (SD) age of this sub-sample
38
39 was 41.9 (13.02) years, 71.6% were female, while 21.0% of patients were unemployed and
40
41 3.8% were non-white. A flowchart describing how the samples of patients (N_P) and
42
43 therapists (N_T) were derived is presented in Figure 1.
44
45
46
47

48 49 *Baseline and outcome variables*

50
51 Baseline patient demographic and severity data were collected using the CORE Assessment
52
53 form (Barkham et al., 2005) and CORE-OM (Barkham et al., 2001; Evans et al., 2002). The
54
55 CORE-OM is a self-report measure of a patient's condition over the past week and
56
57 comprises 34 items addressing the domains of subjective wellbeing, symptoms, functioning,
58
59
60

1
2
3 and risk. The risk domain captured both risk-to-self (4 items: e.g., I have made plans to end
4 my life) and risk-to-others (2 items: e.g., I have been physically violent to others). Items are
5 scored on a 0 to 4 scale and yield an overall CORE-OM score that can be separated into a
6 CORE non-risk score and CORE risk score, each with a range from 0 to 40. The 34-item scale
7 has a reported internal consistency of .94 (Barkham et al., 2001) and a one-month test-
8 retest correlation of .88 (Barkham, Mullin, Leach, Stiles, & Lucock, 2007).
9

10
11
12
13
14
15
16
17 Patients completed the CORE-OM prior to therapy and at the end of their final treatment
18 session. Therefore final outcome scores were not available for patients that dropped out of
19 therapy. The two outcomes for the study were whether patients had completed or dropped
20 out of therapy, as recorded by the therapist at case closure and whether those patients that
21 completed therapy deteriorated or not as reflected in their CORE-OM score.
22
23

24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
Reliable change in CORE-OM scores has been defined as a pre-post change in CORE-OM
scores of five points or more (Connell et al., 2007). However, for the reasons stated above
and due to the rarity of reliable deteriorations, pre-post deteriorations of fewer than five
points were also considered.

Analysis

Subsequent to describing patient intake severity and patient outcomes, MLM was
used to produce a multilevel model for each outcome. MLM is a recommended method
where there is a hierarchical structure in the data (i.e., where patients at level 1 are 'nested'
within therapists at level 2) and differences between the higher-level units (i.e., therapists)
are of interest (Goldstein & Spiegelhalter, 1996; Snijders & Bosker, 2012). Explanatory
variables were added to the models, with continuous variables grand mean centred
(Hofmann & Gavin, 1998) and tested for significance by dividing the derived coefficients by
their standard errors. Values greater than 1.96 were considered significant at the 5% level.

1
2
3 Multilevel modeling software, MLwiN v2.30 (Rasbash, Charlton, Browne, Healy, &
4
5 Cameron, 2009) was used to estimate the parameters in each model, initially by marginal
6
7 quasi-likelihood (MQL) methods, before applying these estimates as 'priors' for Markov
8
9 chain Monte Carlo (MCMC) estimation procedures. This simulation approach produces a
10
11 large number of estimates of the unknown parameters that can be summarised to both a
12
13 mean estimate and a 50th percentile estimate. In addition, a 95% probability interval (Pri),
14
15 analogous to 95% confidence intervals, can be taken as the 2.5 and 97.5 percentile values
16
17 (Browne, 2012). During development, MCMC models were compared using the Deviance
18
19 Information Criteria (DIC), which balances 'fit' and 'complexity', with reductions in DIC
20
21 indicating improvements in the model fit (Spiegelhalter, Best, Carlin, & van der Linde, 2002).
22
23

24
25 Because the study samples used in these analyses are much reduced compared with the
26
27 full dataset, sensitivity analysis was carried out. Logistic regression models were developed
28
29 for larger data samples, where exclusion criteria were not applied, and the included
30
31 predictor variables and their odds ratios (ORs) were compared with those derived from the
32
33 smaller study samples.
34
35
36

37
38 The therapist effect on outcome is defined as the percentage of the total variance that is
39
40 at level 2 (therapist level). In the current study, variance on the logistic scale derived from a
41
42 linear threshold method was used (Rasbash et al., 2009; Snijders, & Bosker, 2012).
43
44 Assumptions of normality in the data were tested by plotting the patient level and therapist
45
46 level residuals produced by the model to normal distribution curves using quantile-quantile
47
48 (q-q) plots.
49
50

51
52 The residual for each therapist represents the degree to which a therapist's outcomes
53
54 depart from those of the average therapist while controlling for patient characteristics
55
56 (case-mix) and can be seen as the additional, unexplained impact of the therapist on
57
58
59
60

1
2
3 outcome (Goldstein & Spiegelhalter, 1996; Rasbash et al., 2009; Saxon & Barkham, 2012).

4
5 The therapist residuals from the dropout and deterioration models were considered
6
7 separately by ranking and plotting with confidence intervals (CIs; Goldstein & Healy, 1995;
8
9 Rasbash et al., 2009). Thus for each outcome, therapists could be described as average,
10
11 where their CI crossed the average (residual = 0) in their impact on outcome, while those
12
13 that did not cross the average were identified as significantly above or below average.
14
15

16
17 The therapist residuals from the two models were also plotted against each other as a
18
19 scatterplot, placing each therapist in one of four quadrants: Quadrant 1 comprising those
20
21 therapists better than average on both outcomes; Quadrant 2 those therapists worse than
22
23 average on both outcomes; and in Quadrants 3 and 4, those therapists better on one and
24
25 worse on the other outcome.
26
27

28 29 **Results**

30
31 The results are presented in three main sections, reflecting the three study aims. The two
32
33 sections on dropout and deterioration begin with descriptives of the samples, followed by
34
35 descriptions of the multilevel models and the reporting of therapist effects. The models and
36
37 significant case-mix variables are presented in Appendix A, Appendix B and Table 1. The
38
39 third section of the Results compares and combines the results found for dropout and
40
41 deterioration.
42
43

44 45 **Patient dropout**

46
47 For the dropout sample (N = 10,521), the proportion of patients who dropped out of
48
49 therapy was 33.8%, with over half of these (52.7%) dropping out before session 3. The mean
50
51 (SD) number of sessions attended for dropouts was 2.8 (1.91) sessions, compared with 6.1
52
53 (2.68) for treatment completers. The mean (SD) patient dropout rate for therapists was
54
55 31.5% (13.8) with a range between 1.2% - 73.2% (IQR: 23.6% - 39.9%).
56
57
58
59
60

1
2
3 The mean (SD) patient CORE-OM score at intake was 18.1 (6.31) with 90.0% of patients
4
5 scoring above the clinical cut-off score of 10. For patients who dropped out of therapy (N=
6
7 3,554), the mean (SD) intake score was 18.9 (6.28) and 91.8% were above clinical cut-off.
8
9 This compares to 17.8 (6.28) and 89.1% for patients who completed therapy (N=6,967).

12 *Dropout model development*

14 A single level logistic regression model containing significant predictors of outcome (drop-
15
16 out or not) was developed, prior to extending it to a multi-level model to allow for therapist
17
18 variability. Following MCMC procedures, the difference between the DICs of the multilevel
19
20 model compared to the single level model (688.7) indicated that the multilevel model was a
21
22 better fit for the data. Tests of convergence showed a chain length of 57,000 iterations to be
23
24 sufficient and q-q plots were fairly linear, indicating that Normality can be assumed. The
25
26 dropout multilevel model is presented in Appendix A.
27
28
29

30
31 Table 1 shows the patient variables identified as predictors of dropout, with their odds
32
33 ratios (ORs) and 95% probability intervals (PrIs) produced by the exponentials of the 2.5, 50
34
35 and 97.5 percentile values for the model coefficients. Patients who were younger, non-
36
37 white, unemployed, or had higher CORE non-risk scores were more likely to drop out.
38
39

40
41 In addition, patients answering in the affirmative (either: *only occasionally, sometimes,*
42
43 *often, or most of the time*) to the risk questions 'I have hurt myself physically or taken
44
45 dangerous risks with my health' (N=850; OR=1.19) and 'over the past week I have been
46
47 physically violent to others' (N= 534; OR= 1.39), were both predictive of dropout compared
48
49 to patients indicating no risk on these items. There were no significant interactions between
50
51 variables in the model. In relation to risk, this suggests that the two questions, 'risk to self'
52
53 and 'risk to others', are identifying two separate types of risk. This is supported by the data
54
55 showing that of those patients reporting risk on either item (N=2,316), only 19% scored on
56
57
58
59
60

1
2
3 both items. No significant random slopes were found, indicating that each of the variables in
4
5 the model impacted on outcomes similarly for all therapists.
6

7
8 Sensitivity analysis was carried out on a sample (N=38,354), representing all patients
9
10 accepted for therapy (N=55,070) minus those with missing data (N=16,715). A single level
11
12 logistic regression model produced by the larger data sample contained the same significant
13
14 variables as above and minimal differences in ORs. The variable showing the greatest
15
16 difference was 'Ethnicity' with an OR (95% PrI) of 1.12 (1.01, 1.23) in the larger sample
17
18 compared with 1.29 (1.05, 1.59) in the study sample.
19
20

21 22 *Therapist effects for dropout*

23
24 Individual therapists had a varying impact on outcome after controlling for the significant
25
26 patient predictors identified above, with a significant therapist effect (95% PrI) of 12.6%
27
28 (9.1, 17.4). No therapist factors were available but number of patients per therapist was
29
30 considered in the model and was found to have minimal effect, reducing the therapist effect
31
32 to 12% but indicating a poorer model fit (larger DIC). Therefore the final treatment dropout
33
34 model (Appendix A) included only patient variables.
35
36
37

38
39 Figure 2 plots the therapist intercept residuals (with 95% CIs) produced by the model for
40
41 the 85 therapists ranked best to worst, from left to right. The plot shows that the majority
42
43 of therapists (61.1%), shown in grey, had treatment ending outcomes that were not
44
45 significantly different to the average therapist (indicated by the dashed horizontal line
46
47 where the residual is zero), while 13 (15.3%) therapists, on the left of the chart, had
48
49 significantly better than average outcomes and 20 (23.5%), on the right of the chart, had
50
51 outcomes that were significantly poorer than average. In order to gauge the actual
52
53 differences in dropout rates between these three groups of therapists, their aggregated
54
55 means were calculated. The aggregated mean (SD) dropout rate for average therapists was
56
57
58
59
60

1
2
3 29.7% (6.4), while for above average therapists it was 12.0% (7.3) compared with 49.0%
4
5 (10.4) for below average therapists.

8 **Patient Deterioration**

9
10 For the deterioration sample (N=6,405), where patients completed therapy, the mean (SD)
11
12 CORE-OM score at intake was 17.8 (6.23) while the proportion scoring above clinical cut-off
13
14 was 89.1%. Their mean (SD) outcome score of 8.9 (6.25) yielded a pre-post effect size of
15
16 1.43. Most patients (72.2%) improved by 5 points or more on the CORE-OM and could be
17
18 considered reliably improved, while 26.8% made no reliable change, 6.2% deteriorated to
19
20 some degree, and 1.0% reliably deteriorated. The mean (SD) reliable deterioration rate for
21
22 therapists was 1.2% (1.67) with a range between 0% and 7.1% (IQR: 0 - 1.9%).
23
24
25
26

27 Table 2 shows the deterioration rates for six different levels of deterioration, ranging
28
29 from any change on the CORE-OM to a change of ≥ 5 CORE-OM points (the degree of change
30
31 considered as reliable deterioration) and the number of therapists that had no
32
33 deteriorations for each level. There were significant positive correlations (one-tailed, all p -
34
35 values < 0.001) between the different rates and rankings for therapists. Correlation
36
37 coefficients (Spearman's ρ) ranged from .50 for the association between 'any
38
39 deterioration' and ' ≥ 5 ' point change, to .92 for the association between 'any deterioration'
40
41 and '>1' point change.
42
43
44

45 The large proportion of therapists with no deteriorations was problematic in multilevel
46
47 model development and only where deterioration was defined as 'any deterioration' or '> 1'
48
49 did the models stabilise to produce reliable estimates of therapist effects. Therefore, a
50
51 model with deterioration at the level of 'more than 1 point' was used as the patient
52
53 outcome in the multilevel analysis. The correlation between therapists ranked using this
54
55 level of deterioration and reliable deterioration (' ≥ 5 ') was .56 ($p < .001$).
56
57
58
59
60

Deterioration model development

As with the dropout model, a single level logistic regression model containing significant predictors of outcome was extended to allow for therapist variability. Following MCMC procedures, the change in DIC of 45.9 indicated the multilevel model to be a better fit for the data than the single level model. Tests of convergence indicated that the chain length of 128,000 iterations was sufficient and the q-q plots were fairly linear, indicating that Normality can be assumed. The deterioration model is presented in Appendix B.

Table 1 shows the patient variables identified as predictors of deterioration by more than 1 point. Patients who were older and less severe at intake were more likely to deteriorate. However, the latter is likely to be a statistical factor with higher scores having less scope to deteriorate. The risk item 'I have thought of hurting myself', was a significant predictor of deterioration (N=1,829; OR= 1.55) and, consistent with the drop-out model, patients who were unemployed were more likely to deteriorate than patients not unemployed. Again, there were no interactions between variables and no significant random slopes on any of the predictor variables indicating that they have a similar impact on outcome for all therapists.

Sensitivity analysis, using the largest possible sample (N = 24,499) representing all those patients who completed therapy and had a pre and post CORE OM score (N = 30,978) minus those with missing variable data (N= 6,479), produced a logistic regression model containing the same four predictor variables as in Table 1. The ORs for age and CORE non-risk score were almost identical to those produced by the smaller sample. The ORs (95%PrI) for unemployment and the risk question, of 2.04 (1.22, 2.33) and 1.41 (1.21, 1.66) respectively,

1
2
3 were reduced, although for both variables the PrIs overlap their corresponding PrIs derived
4
5 from the smaller samples.
6

7 *Therapist effects for deterioration*

8
9
10 The therapist effect for deterioration of more than 1 point was 10.1% (95% PrI: 4.9, 17.8).

11
12 The number of patients per therapist was not significant in the model. As with the
13
14 caterpillar plot for dropout (Figure 2), Figure 3 plots the therapist intercept residuals
15
16 produced by the deterioration model (with 95% CIs) for the 85 therapists ranked best to
17
18 worst, from left to right.
19
20

21
22 Indicative of the rarity of the event and the smaller numbers of patients per therapist,
23
24 the 95% CIs are generally wider than in the dropout model, with only one therapist being
25
26 significantly better than average, and four therapists significantly worse than average. The
27
28 vast majority of therapists (94.1%) could be considered average with regard to patient
29
30 deterioration, they had an overall mean (SD) deterioration rate of 4.6% (3.7). The better
31
32 than average therapist had no patients who deteriorated, while for the four below average
33
34 therapists, their rates of deterioration were, from left to right, 11.8%, 12.1%, 14.1% and
35
36 14.9%. The statistically reliable deterioration rates (deterioration by ≥ 5 points) for these
37
38 four therapists, were 1.5%, 3.5%, 3.1% and 3.0% respectively, compared with a mean (SD)
39
40 rate of 1.1% (1.7), for the average therapists.
41
42
43
44

45 **Combining therapist variability on dropout and deterioration**

46
47
48 In order to consider whether those therapists with more treatment dropouts also had more
49
50 treatment completers that deteriorated, the therapist rankings and residuals from Figures 2
51
52 and 3 were compared. There was no significant correlation between the rankings
53
54 (Spearman's $\rho = 0.07$, $p=0.52$), suggesting that, overall, therapists that were less able to
55
56 retain patients in therapy did not generally have more patients that deteriorated after
57
58
59
60

1
2
3 completing treatment. To consider the relationship between the two outcomes further, the
4 therapist residuals for each outcome were plotted against each other in a scatterplot (Figure
5
6
7
8 4).

9
10 In Figure 4, the x-axis measures the therapist residual for dropout, while the y-axis
11
12 measures the therapist residual for deterioration. Zero on each axis represents the average
13
14 therapist and each therapist is placed in a quadrant of the plot based on their residuals
15
16 derived from each model. The 20 therapists significantly below average for dropout are
17
18 represented by black circles while the four therapists identified as significantly below
19
20 average for deterioration are represented by grey squares. The 95% CIs from Figures 2 and
21
22 3, which would be represented by a cross through every therapist point, are not shown, but
23
24 in all instances at least one CI crossed zero. Therefore, no therapist was found to be
25
26 significantly below average on both outcomes.
27
28
29
30

31 Discussion

32
33 In this practice-based study, our aim was to establish the degree to which therapists
34
35 contribute to the variability in two negative patient outcomes, namely unplanned endings
36
37 (i.e., dropouts) and deterioration. For both outcomes, we found significant therapist effects,
38
39 of 12.6% and 10.1% respectively, that were larger than the range of effects (5%-8%) found
40
41 in similar studies of patient improvement (e.g., Lutz, Leon, Martinovich, Lyons, & Stiles,
42
43 2007; Wampold & Brown, 2005). In a context where the overall effect of therapy, which
44
45 includes all aspects of therapy including therapist factors, treatment adherence, and alliance
46
47 is estimated at 20% (Baldwin & Imel, 2013), these therapist effects of over 10% are both
48
49 statistically significant and clinically important.
50
51
52
53

54
55 Locating the focus for patient outcomes with the therapist supports findings from studies
56
57 of addiction services (Brorson, Arnevik, Rand-Hendriksen, & Duckert, 2013) and adolescent
58
59
60

1
2
3 services (de Haan, Boon, de Jong, Hoeve, & Vermeiren, 2013). These studies concluded that
4
5 the simple study of patient variables in isolation was of limited value and the study of such
6
7 factors as the alliance and therapist variables would be more useful, in part because they
8
9 are variables that can be changed (de Haan et al., 2013).
10

11
12 Therapist variables that have been associated with negative outcome include lack of
13
14 empathy, negative countertransference, overuse of transference interpretations, and
15
16 disagreement with patients about therapy process (Mohr, 1995). Type and amount of
17
18 training, theoretical orientation, and gender were not predictive of patient outcome (Okiishi
19
20 et al., 2006), while studies of therapist competence, have yielded contradictory results
21
22 (Ginzburg et al., 2012; Webb, de Rubeis & Barber, 2010). Our finding that those therapists
23
24 worse than average for dropout were no worse than average for deterioration, suggests
25
26 that different therapist factors may be associated with different negative outcomes. Further
27
28 research is necessary to identify therapist factors and their interactions with patient
29
30 characteristics that may explain the degree of variability between therapists in their
31
32 negative outcomes.
33
34
35
36
37
38

39 Our finding that around a third of patients dropped out of therapy is within the range
40
41 of 20%-60% reported elsewhere (Reneses et al., 2009) and is 10% larger than reported
42
43 rates where session 1 was excluded (e.g., RCP, 2011; 2013). The mean therapist rate of
44
45 31.5% was similar to the 33.2% found by Zimmerman et al (2016), however, our therapist
46
47 effect for dropout was twice that found in their study. We can only speculate as to why
48
49 there was such a difference, but reasons may include differences in methodology
50
51 (Goldstein, Rasbash & Browne, 2002; Snijders & Bosker, 2012), sample size (Soldz, 2006),
52
53 service delivery models and available patient variables.
54
55
56
57
58
59
60

1
2
3 Our analysis identified 23.5% of therapists whose dropout rates were significantly
4 higher than average. Aggregated dropout rates indicated that patients seen by these
5 therapists were around four times more likely to dropout than patients seen by the 15.3% of
6 therapists who had significantly lower than average dropout (49% compared with 12%).
7
8
9

10
11 The results for patient deterioration were less reliable, reflected in the wide
12 Probability Interval for the therapist effect and the wider confidence intervals for therapist
13 residual estimates. This unreliability was due to the rarity of the outcome, the smaller
14 number of patients per therapist and the adoption of a measure of deterioration that was
15 less than 'statistically reliable'. That said, where patient safety and possible harm are
16 paramount, it would seem appropriate to 'flag' therapists at the below average end, as soon
17 as possible, regardless of the confidence intervals or number of patients they have treated.
18
19 We found significant outcome variability between therapists, with patients seen by
20 therapists identified as below average being over twice as likely to deteriorate as patients
21 seen by therapists identified as average. That those therapists identified as below average,
22 using our less stringent criteria, also showed higher than average rates of *reliable*
23 deterioration suggests that the model is correctly identifying therapists with higher rates of
24 negative change.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42

43 *Case-mix variables*

44
45
46 A number of patient variables were significant predictors of outcomes and were
47 controlled for in estimating the impact of the therapist. We found that these variables
48 effected therapists similarly, i.e., there were no random slopes. For dropout, the patient
49 variables identified were similar to those reported elsewhere: greater symptom severity at
50 intake (Kazdin, Mazurick, & Siegel, 1994); younger age (e.g., Edlund et al., 2002), and non-
51
52
53
54
55
56
57
58
59
60

1
2
3 white ethnicity and unemployment, which may be proxy measures of socio-economic
4
5 deprivation (Garfield, 1994; Wierzbiki & Pekarik, 1993; Williams, Ketring, & Salts, 2005). In
6
7 addition, and possibly of greater concern, was the finding that patients at risk of harming
8
9 themselves or others were more likely to dropout than patients with no risk, a finding that
10
11 supports previous research from a single service study using CORE risk items (Saxon,
12
13 Ricketts & Heywood, 2010). We found that patients who had been 'physically violent to
14
15 others' were 39% more likely to dropout than those who had not.
16
17
18
19

20 For deterioration, we found that in addition to answering in the affirmative to the risk
21
22 question 'over the past week I have thought of hurting myself', patient age and employment
23
24 status were also predictive of outcome. Younger patients were more likely to drop out than
25
26 older patients, but if they completed therapy they were less likely to have deteriorated,
27
28 while unemployed patients were 44% more likely to drop out than patients who were not
29
30 unemployed, and if they stayed in therapy to a planned ending they were more than twice
31
32 as likely to have deteriorated.
33
34
35
36

37 *Study Limitations*

38
39 Crucial in any practice-based study is the issue of the representativeness of included data
40
41 (Brown, Lambert, Jones, & Minami, 2005). In order to reduce any bias due to the failure to
42
43 collect data from patients, only those therapists who provided data for over 90% of their
44
45 patients were included, therefore results may only be generalizable to therapists with high
46
47 return rates. Also, our sample contained counsellors and clinical psychologists in primary
48
49 care who had seen at least 30 patients for dropout or 13 for deterioration, therefore results
50
51 may not be generalizable to therapists who have seen fewer patients or deliver other types
52
53
54
55
56
57
58
59
60

1
2
3 of therapy in different settings. The small number of sites, and therapists per site,
4
5 prevented any analysis of the impact treatment sites might have on both outcomes.
6
7

8 In addition to concerns about the reliability of the deterioration analysis outlined above,
9
10 the deterioration rates reported here may underestimate actual rates as they are based on
11
12 treatment completers only. No last CORE-OM was available for patients who dropped out,
13
14 therefore it was not possible to measure their pre-post change, but research indicates that
15
16 they are likely to have had poorer clinical outcomes (Delgado et al., 2014; Saxon, Firth &
17
18 Barkham, submitted 2016). To address these limitations, it would be informative to
19
20 replicate this analysis with a larger multi-site dataset that contains a wider range of patient
21
22 and therapist variables and outcome measures for the last session attended.
23
24
25
26

27 *Clinical and service implications*

28
29
30 These results have important implications for quality improvement in psychological therapy
31
32 services. Services may not be meeting the needs of some sections of the community and
33
34 should take steps to better engage patients who are younger, of non-white ethnicity or
35
36 unemployed. With regard to risk, heightened patient risk may be associated with greater
37
38 severity and complexity of condition and possible borderline personality disorder.
39

40
41 Guidelines suggest that brief, psychological therapies in primary care are unsuitable for
42
43 patients with borderline personality disorder, who have higher levels of self-harm, or anti-
44
45 social personality disorder where higher levels of aggression are characteristic, therefore
46
47 patients may need to be referred-on to more appropriate services (NICE, 2009a,b).
48

49
50
51 Our results show that patient characteristics alone cannot account for drop out and clinical
52
53 deterioration and that therapists account for a large proportion of the variance in these
54
55 negative outcomes. This is an important factor that is often neglected, for example when
56
57
58
59
60

1
2
3 considering ways of reducing early withdrawal from treatment (Barrett et al. 2008). The
4
5 implication is that therapists who are below average for negative outcomes should be made
6
7 aware of this so that remedial action, for example through greater support, supervision or
8
9 training, can be taken. However, caution is necessary because although the statistical
10
11 methods employed in this study can raise questions about therapist outliers, other
12
13 unmeasured factors may influence therapist performance. Therapists and service managers
14
15 need to use these methods only as a starting point for exploration.
16
17
18
19

20 *Conclusion*

21
22 In conclusion, using sophisticated and appropriate methods, we found large therapist
23
24 effects for both types of negative outcomes, indicating significant variability between
25
26 therapists in their ability to retain patients in therapy and to prevent patient deterioration.
27
28 This study illustrates that the reporting of simple aggregated outcomes for services and
29
30 practitioners, usually focused on improvement and recovery, is limited and may mask
31
32 important factors for safe and effective service delivery in the psychological therapies.
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Acknowledgements

We thank CORE IMS for the collection of the data and in particular John Mellor-Clark and Alex Curtis Jenkins.

Conflict of interest

MB was a member of the research group that developed the CORE-OM.

Funding source

The research was funded by the National Institute for Health Research (NIHR) under its Research for Patient Benefit (RfPB) Programme (Grant Reference Number PB-PG-0408-15144). They had no involvement in any aspect of this manuscript.

References

- 1
2
3
4
5 Baldwin, S.A., & Imel, Z.E. (2013). Therapist effects: Findings and methods. In M. J. Lambert
6
7 (Ed.) *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed.) (pp.
8
9 258-297). Hoboken, NJ: Wiley.
- 10
11
12 Barkham, M., Gilbert, N., Connell, J., Marshall, C., & Twigg, E. (2005). Suitability and utility of
13
14 the CORE-OM and CORE-A for assessing severity of presenting problems in psychological
15
16 therapy services based in primary and secondary care settings. *British Journal of*
17
18 *Psychiatry*, *186*, 239-246.
- 19
20
21 Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans C.,... & McGrath G.
22
23 (2001). Service profiling and outcomes benchmarking using the CORE-OM: Towards
24
25 practice-based evidence in the psychological therapies. *Journal of Consulting & Clinical*
26
27 *Psychology*, *69*, 184-196.
- 28
29
30
31 Barkham, M., Mullin, T., Leach, C., Stiles, W.B., & Lucock, M. (2007). Stability of the CORE-
32
33 OM and BDI-I: Psychometric properties and implications for routine practice. *Psychology*
34
35 *& Psychotherapy: Theory, Research & Practice*, *80*, 269-278.
- 36
37
38 Barrett, M., Chua, W., Crits-Christoph, P., Gibbons, M., & Thompson, D. (2008). Early
39
40 withdrawal from mental health treatment: Implications for psychotherapy practice.
41
42 *Psychotherapy: Theory, Research, Practice, Training*, *45*, 247-267.
- 43
44
45 Brorson, H.H., Arnevik, E.A., Rand-Hendriksen, K., & Duckert, F. (2013). Drop-out from
46
47 addiction treatment: A systematic review of risk factors. *Clinical Psychology Review*, *33*,
48
49 1010-1024.
- 50
51
52 Brown, G.S., Lambert, M.J., Jones, E.R., & Minami, T. (2005). Identifying highly effective
53
54 psychotherapists in a managed care environment. *American Journal of Managed Care*.
55
56 *11*, 513-520.
57
58
59
60

- 1
2
3 Browne, W.J. (2012). *MCMC Estimation in MLwiN*, v2.26. Centre for Multilevel Modelling,
4
5 University of Bristol.
6
7
8 Cahill, J., Barkham, M., Hardy, G., Rees, A., Shapiro, D.A., Stiles, W.B., & Macaskill, N. (2003).
9
10 Outcomes of patients completing and not completing cognitive therapy for depression.
11
12 *British Journal of Clinical Psychology*, 42, 133–143.
13
14
15 Cahill, J., Barkham, M., & Stiles, W.B. (2010). Systematic review of practice-based research
16
17 on psychological therapies in routine clinic settings. *British Journal of Clinical Psychology*,
18
19 49, 421-454.
20
21
22 Castonguay, L.G., Barkham, M., Lutz, W., & McAleavy, A. (2013). Practice oriented research:
23
24 Approaches and applications. In M.J. Lambert (Ed.), *Bergin & Garfield's handbook of*
25
26 *psychotherapy and behavior change*. 6th Edition. Hoboken, N.J.: Wiley. pp. 85-133.
27
28
29
30
31 Connell, J., Barkham, M., Stiles, W.B., Twigg, E., Singleton, N., Evans, O., & Miles, J.N.V.
32
33 (2007). Distribution of CORE-OM scores in a general population, clinical cut-off points,
34
35 and comparison with the CIS-R. *British Journal of Psychiatry*, 190, 69-74.
36
37
38 Craze, L., McGeorge, P., Holmes, D., Bernardi, S., Taylor, P., Morris-Yates, A., & McDonald, E.
39
40 (2014). *Recognising and responding to deterioration in mental state: A scoping review*.
41
42 Sydney: Australian Commission on Safety and Quality in Health Care.
43
44
45 Delgadoillo, J., McMillan, D., Lucock, M., Leach, C., Ali, S., & Gilbody, S. (2014). Early changes,
46
47 attrition, and dose-response in low intensity psychological interventions. *British Journal*
48
49 *of Clinical Psychology* 53(1), 114-130. doi: 10.1111/bjc.12031.
50
51
52 Department of Health (2012). *IAPT three-year report: The first million patients*. Crown
53
54 Copyright. www.dh.gsi.gov.uk Accessed 5th February 2015
55
56
57 de Haan, A.M., Boon, A.E., de Jong, J.T.V.M., Hoeve, M., & Vermeiren, R.R.J.M. (2013). A
58
59
60

1
2
3 meta-analytic review of treatment dropout in child and adolescent outpatient mental
4
5 health care. *Clinical Psychology Review*, 33, 698-711.

6
7 Edlund, M.J., Wang, P.S., Berglund, P.A., Katz, S.J., Lin, E., & Kessler, R. (2002). Dropping out
8
9 of mental health treatment: Patterns and predictors among epidemiological survey
10
11 respondents in the United States and Ontario. *American Journal of Psychiatry*, 159, 845–
12
13 851.

14
15 Evans, C., Connell, J., Barkham, M., Margison, F., Mellor-Clark, J., McGrath, G., & Audin, K.
16
17 (2002). Towards a standardised brief outcome measure: Psychometric properties and
18
19 utility of the CORE-OM. *British Journal of Psychiatry*, 180, 51-60.

20
21 Garfield, S.L. (1994). Research on client variables in psychotherapy. In A.E. Bergin & S.L.
22
23 Garfield (Eds.), *Handbook of psychotherapy and behavior change*. pp. 190-228. Wiley:
24
25 New York.

26
27 Gibbons, C.J., Fournier, J.C., Stirman, S.W., DeRubeis, R.J., Crits-Christoph, P., & Beck, A.T.
28
29 (2010). The clinical effectiveness of cognitive therapy for depression in an outpatient
30
31 clinic. *Journal of Affective Disorders*, 125, 169–176.

32
33 Ginzburg, D. M., Bohn, C., Höfling, V., Weck, F., Clark, D. M., & Stangier, U. (2012).
34
35 Treatment specific competence predicts outcome in cognitive therapy for social anxiety
36
37 disorder. *Behaviour research and therapy*, 50(12), 747-752.

38
39 Glover, G., Webb, M., & Evison, F. (2010). *Improving Access to Psychological Therapies: A*
40
41 *review of the progress made by sites in the first rollout year*. Retrieved June 2013, from
42
43 [http://www.iapt.nhs.uk/silo/files/iapt-a-review-of-the-progress-made-by-sites-in-the-](http://www.iapt.nhs.uk/silo/files/iapt-a-review-of-the-progress-made-by-sites-in-the-first-rollout-year.pdf)
44
45 [first-rollout-year.pdf](http://www.iapt.nhs.uk/silo/files/iapt-a-review-of-the-progress-made-by-sites-in-the-first-rollout-year.pdf).

46
47 Goldstein, H., & Healy M.J.R. (1995). The graphical presentation of a collection of means.
48
49 *Journal of the Royal Statistical Society*, 158, 175-177.

1
2
3 Goldstein, H., Rasbash, J. & Browne, W.J. (2002). Partitioning variation in multilevel models.

4
5 *Understanding Statistics, 1*, 223-231

6
7 Goldstein, H., & Spiegelhalter, D. (1996). League tables and their limitations: statistical

8
9 issues in comparisons of institutional performance-with discussion. *Journal of the Royal*

10
11 *Statistical Society, 159*, 385-443.

12
13
14 Hansen, N., Lambert, M., & Forman, E. (2002). The psychotherapy dose-response effect and

15
16 its implications for treatment delivery services. *Clinical Psychology: Science & Practice, 9*,

17
18 329–343.

19
20 Hofmann, D.A. & Gavin, M.B. (1998). Centering decisions in hierarchical linear models:

21
22 Implications for research in organisations. *Journal of Management, 24*, 623-641

23
24 Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining

25
26 meaningful change in psychotherapy research. *Journal of Consulting and Clinical*

27
28 *Psychology, 59*, 12-19.

29
30 Kazdin, A.E., Mazurick, J.L., & Siegel, T.C. (1994). Treatment outcome among children with

31
32 externalizing disorder who terminate prematurely versus those who complete

33
34 psychotherapy. *Journal of the American Academy of Child & Adolescent Psychiatry, 33*,

35
36 549-557.

37
38 Kraus, D. R., Castonguay, L., Boswell, J. F., Nordberg, S. S., & Hayes, J. A. (2011). Therapist

39
40 effectiveness: Implications for accountability and patient care. *Psychotherapy*

41
42 *Research, 21*(3), 267-276.

43
44 Lambert, M. J. (2010). *Prevention of treatment failure: The use of measuring, monitoring,*

45
46 *and feedback in clinical practice.* Washington D.C.: American Psychological Association.

47
48 Lambert, M. J. (2013). The efficacy and effectiveness of psychotherapy. In M. J. Lambert

49
50 (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp.

- 1
2
3 169-218). New York, NY: Wiley.
4
5 Linden, M. (2013). How to define, find and classify side effects in psychotherapy: From
6
7 unwanted events to adverse treatment reactions. *Clinical Psychology & Psychotherapy*,
8
9 20, 286–296.
10
11 Lutz, W., Leon, S.C., Martinovich, Z., Lyons, J.S., & Stiles, W.B. (2007). Therapist effects in
12
13 outpatient psychotherapy: A three-level growth curve approach. *Journal of Counseling*
14
15 *Psychology*, 54, 32–39.
16
17 Maas, C.J.M., & Hox, J.J. (2005). Sufficient Sample sizes for multilevel modeling.
18
19 *Methodology*, 1(3), 86-92. DOI 10.1027/1614-1881.1.3.86
20
21
22
23
24 Mohr, D. C. (1995). Negative outcome in psychotherapy: A critical review. *Clinical*
25
26 *psychology: Science and practice*, 2(1), 1-27.
27
28
29 National Institute for Health and Care Excellence (2009a) Antisocial personality disorder:
30
31 prevention and management. NICE guideline CG77.
32
33 National Institute for Health and Care Excellence (2009b) Borderline personality disorder:
34
35 recognition and management. NICE guideline CG78.
36
37
38 Okiishi, J.C., Lambert, M.J., Eggett, D., Nielson, S.L., Vermeersch, D.A., & Dayton, D.D. (2006).
39
40 An analysis of therapist treatment effects: Toward providing feedback to individual
41
42 therapists on their patients' psychotherapy outcome. *Journal of Clinical Psychology*, 62,
43
44 1157-1172.
45
46
47 Rasbash, J., Charlton, C., Browne, W.J., Healy, M., & Cameron, B. (2009). *MLwiN Version*
48
49 2.1. Centre for Multilevel Modelling, University of Bristol.
50
51
52 Reneses, B., Munoz, E., & Lopez-Ibor, J.J. (2009). Factors predicting drop-out in community
53
54 mental health centres. *World Psychiatry*, 8, 173-177.
55
56
57 Rogers, C.P. (1951). *Client-centered therapy*. Boston: Houghton Mifflin.
58
59
60

1
2
3 Royal College of Psychiatrists. (2011). *National Audit of Psychological Therapies for Anxiety*
4
5 *and Depression*, National Report.

6
7 Royal College of Psychiatrists. (2013). *Report of the Second Round of the National Audit of*
8
9 *Psychological Therapies (NAPT) 2013*. London: Healthcare Quality Improvement
10
11 Partnership.

12
13 Saxon, D., Ricketts, T., & Heywood, J. (2010). Who drops-out? Do measures of risk to self
14
15 and to others predict unplanned endings in primary care counselling? *Counselling and*
16
17 *Psychotherapy Research*, 10(1), 13-21.

18
19 Saxon, D., & Barkham, M. (2012). Patterns of therapist variability: Therapist effects and the
20
21 contribution of patient severity and risk. *Journal of Consulting & Clinical Psychology*, 80,
22
23 535-546.

24
25 Shepherd, M., Evans, C., Cobb, S., & Ghossain, D. (2012). Does therapy makes things worse?
26
27 Investigating episodes of psychological therapy where clients' scores showed reliable
28
29 deterioration. *Clinical Psychology Forum*, 233, 8-12.

30
31 Snijders, T.A.B., & Bosker, R.J. (2012). *Multilevel analysis: An introduction to basic and*
32
33 *advanced multilevel modeling*, (2nd edition). London: Sage Publishers.

34
35 Soldz, S. (2006). Models and meanings: Therapist effects and the stories we tell.
36
37 *Psychotherapy Research*, 16, 173-177.

38
39 Spiegelhalter, D.J., Best, N.G., Carlin, B.P., & van der Linde A. (2002). Bayesian measures of
40
41 model complexity and fit. *Journal of the Royal Statistical Society. Series B (Statistical*
42
43 *Methodology)*, 64, 583-639.

44
45 Stiles, W.B., Barkham, M., Connell, J., & Mellor-Clark, J. (2008). Responsive regulation of
46
47 treatment duration in routine practice in United Kingdom primary care settings:
48
49 Replication in a larger sample. *Journal of Consulting & Clinical Psychology*, 76, 298-305.
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 Swift, J.K., & Greenberg, R.P. (2012). Premature discontinuation in adult psychotherapy: a
4
5 meta-analysis. *Journal of Consulting & Clinical Psychology, 80*, 547-559.
6
7
8 Wampold, B.E., & Brown, G.S. (2005). Estimating variability in outcomes attributable to
9
10 therapists: a naturalistic study of outcomes in managed care. *Journal of Consulting &*
11
12 *Clinical Psychology, 73*, 914-923.
13
14
15 Webb, C.A., DeRubeis, R.J., & Barber, J.P. (2010). Therapist adherence/competence and
16
17 treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical*
18
19 *Psychology, 78*(2), 200-211. <http://dx.doi.org/10.1037/a0018912>
20
21
22 Westmacott, R., Hunsley, J., Best, M., Rumstein-McKean, O., & Schindler, D. (2010). Client
23
24 and therapist views of contextual factors related to termination from psychotherapy: A
25
26 comparison between unilateral and mutual terminators. *Psychotherapy Research, 20*,
27
28 423-435.
29
30
31 Wierzbicki, M., & Pekarik, G. (1993). A meta-analysis of psychotherapy dropout. *Professional*
32
33 *Psychology: Research & Practice, 24*, 190-195.
34
35
36 Williams, S.L, Ketring, S.A., & Salts, C.J. (2005). Premature termination as a function of intake
37
38 data based on ethnicity, gender, socioeconomic status, and income. *Contemporary*
39
40 *Family Therapy, 27*, 213-231.
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1: Odds Ratios for the predictor variables in each model, with their 95% Probability Intervals (PrIs)

Variable in model	Odds Ratios (95% PrI)	
	Drop-out Model	Deterioration Model
Unemployed	1.44 (1.30, 1.60)	2.71 (2.05, 3.57)
Age - grand mean	0.97 (0.96, 0.97)	1.02, (1.01, 1.03)
CORE non-risk – grand mean	1.02 (1.01, 1.02)	0.90 (0.88, 0.92)
Ethnicity (not white)	1.29 (1.05, 1.59)	NS
'I have been physically violent to others'	1.39 (1.21, 1.60)	NS
'I have hurt myself physically or taken dangerous risks with my health'	1.19 (1.05, 1.34)	NS
'I have thought of hurting myself'	NS	1.55 (1.12, 2.14)

Table 2: The number of patients who deteriorated for each level of deterioration, ranging from any deterioration to deterioration of 5 points or more, the mean (SD) deterioration rates for therapists (N=85) and the number of therapists with no deteriorations at each level.

Deterioration	Patient rate N (%)	Therapist rate Mean (SD)	Therapist rate Range %	N (%) of therapists with no deteriorations
Any deterioration	399 (6.2)	6.8 (5.29)	0 – 28.6	11 (12.9)
>1 point	287 (4.5)	5.0 (4.12)	0 – 15.4	16 (18.8)
>2 points	191 (3.0)	3.2 (2.87)	0 – 10.3	24 (28.2)
> 3 points	134 (2.1)	2.2 (2.17)	0 – 7.7	31(36.5)
> 4 points	93 (1.5)	1.7 (1.96)	0 – 7.7	36 (42.4)
5 or more points	67 (1.0)	1.2 (1.67)	0 – 7.1	44 (51.8)

$$\begin{aligned} \text{Ending}_{ij} &\sim \text{Binomial}(\text{denom}_{ij}, \pi_{ij}) \\ \text{logit}(\pi_{ij}) &= \beta_{0j}\text{cons} + -0.030(0.002)(\text{Age-gm})_{ij} + 0.365(0.053)\text{EmployStatus_1}_{ij} + \\ &\quad 0.259(0.106)\text{Ethnicity_1}_{ij} + 0.015(0.003)(\text{CoreXRpre-gm})_{ij} + 0.331(0.071)\text{coR06p_1}_{ij} + \\ &\quad 0.170(0.063)\text{coR34p_1}_{ij} \\ \beta_{0j} &= -1.043(0.084) + u_{0j} \\ [u_{0j}] &\sim N(0, \Omega_u) : \Omega_u = [0.483(0.094)] \\ \text{var}(\text{Ending}_{ij} | \pi_{ij}) &= \pi_{ij}(1 - \pi_{ij}) / \text{denom}_{ij} \\ \text{Deviance(MCMC)} &= 12187.190(10521 \text{ of } 10521 \text{ cases in use}) \end{aligned}$$

Appendix A: Multilevel model for patient dropout. Key: **Ending_{ij}**: Outcome (drop-out or not) for patient *i* seen by therapist *j*. Outcomes follow a **Binomial** distribution with parameters, **denom_{ij}** (=1 for binary outcomes) and π_{ij} (the probability that patient *i* seen by therapist *j* will drop-out).

logit(π_{ij}): the link function used (logit).

$\beta_{0j}\text{cons}$: The intercept for therapist *j*.

The coefficients (SE) of included patient variables: **Age-gm** (age minus grand mean); **EmployStatus_1** (unemployed); **Ethnicity_1** (not white); **CoreXRpre-gm** (pre CORE score excluding risk, minus grand mean); **coR06p_1** (affirmative response to CORE question 6, 'over the past week I have been physically violent to others'); **coR34p_1** (affirmative response to CORE question 34, 'I have hurt myself physically or taken dangerous risks with my health').

β_{0j} : Therapist intercept consisting of two terms, the fixed component (SE) = -1.043 (0.084) plus the therapist specific component, **u_{0j}** .

The random effect [**u_{0j}**] follows a Normal distribution (**N**) with a mean of 0 and a variance (Ω_u) of 0.483 (SE:0.094).

Var(Ending_{ij} | π_{ij}) = $\pi_{ij}(1 - \pi_{ij}) / \text{denom}_{ij}$: states that the variance of the binomial outcome = the probability that patient *i* seen by therapist *j* will drop-out, multiplied by, 1 minus the probability that patient *i* seen by therapist *j* will drop-out (denom=1).

Deviance (MCMC): used to produce the Deviance Information Criteria, an indicator of model 'fit'.

157x87mm (132 x 99 DPI)

$$\begin{aligned}
 & \text{Deterioration} > 1_{ij} \sim \text{Binomial}(\text{denom}_{ij}, \pi_{ij}) \\
 & \text{logit}(\pi_{ij}) = \beta_{0j} \text{cons} + -0.104(0.010)(\text{CoreXRpre-gm})_{ij} + 0.995(0.143)\text{EmployStatus}_1_{ij} + \\
 & \quad 0.440(0.166)\text{coR09p}_1_{ij} + 0.022(0.005)(\text{Age-gm})_{ij} \\
 & \beta_{0j} = -3.782(0.134) + u_{0j} \\
 & [u_{0j}] \sim N(0, \Omega_u) : \Omega_u = [0.389(0.140)] \\
 & \text{var}(\text{Deterioration} > 1_{ij} | \pi_{ij}) = \pi_{ij}(1 - \pi_{ij}) / \text{denom}_{ij} \\
 & \text{Deviance}(MCMC) = 2095.701(6389 \text{ of } 6405 \text{ cases in use})
 \end{aligned}$$

Appendix B: Multilevel model for deterioration of more than one point on CORE-OM. Key: **Deterioration** > **1_{ij}** : Outcome (deterioration by more than 1 point on CORE-OM or not) for patient **i** seen by therapist **j**. Patient variables **CoreXRpre-gm**, **EmployStatus_1** and **Age-gm** are described in Appendix 1, **coR09p_1** indicates an affirmative response to CORE question 9, 'I have thought of hurting myself'. The model notation and other parameters are described in Appendix A.

177x91mm (96 x 96 DPI)

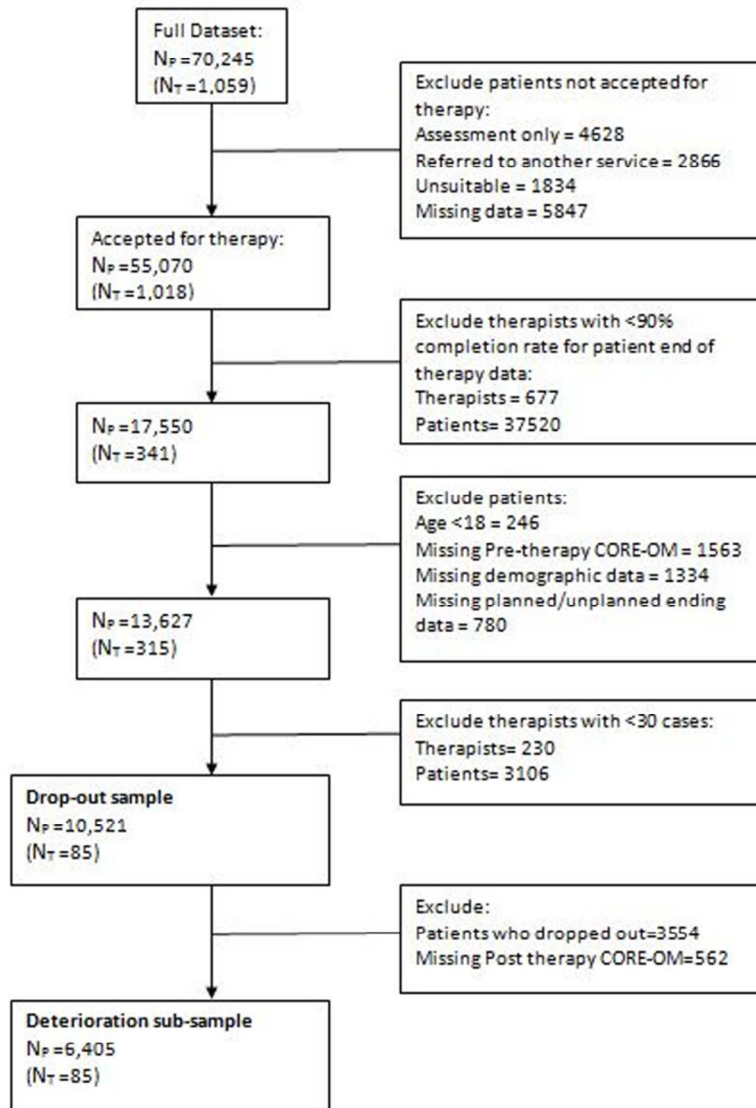


Figure 1: Flowchart showing how the study samples were derived from the full data sample.
203x266mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

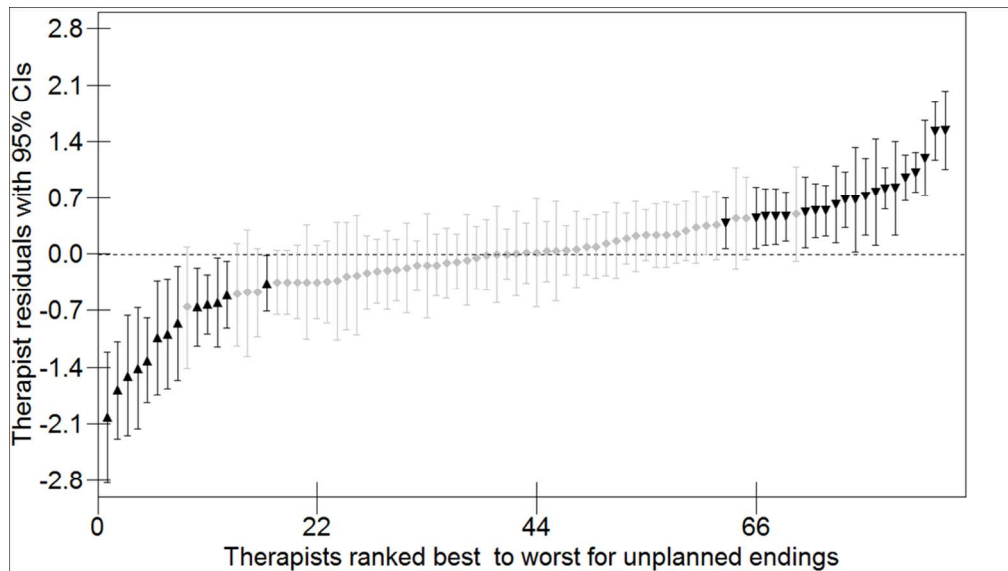


Figure 2: Plot of therapist residuals (with 95% CI) for patient dropout.
226x128mm (132 x 132 DPI)

Peer Review

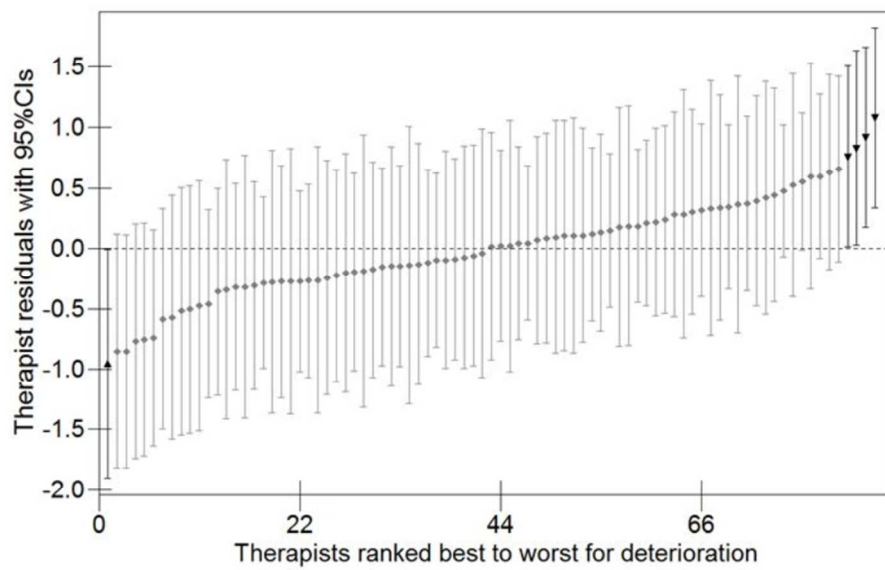


Figure 3: Plot of therapist residuals (with 90% CIs) for patient deterioration
270x172mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

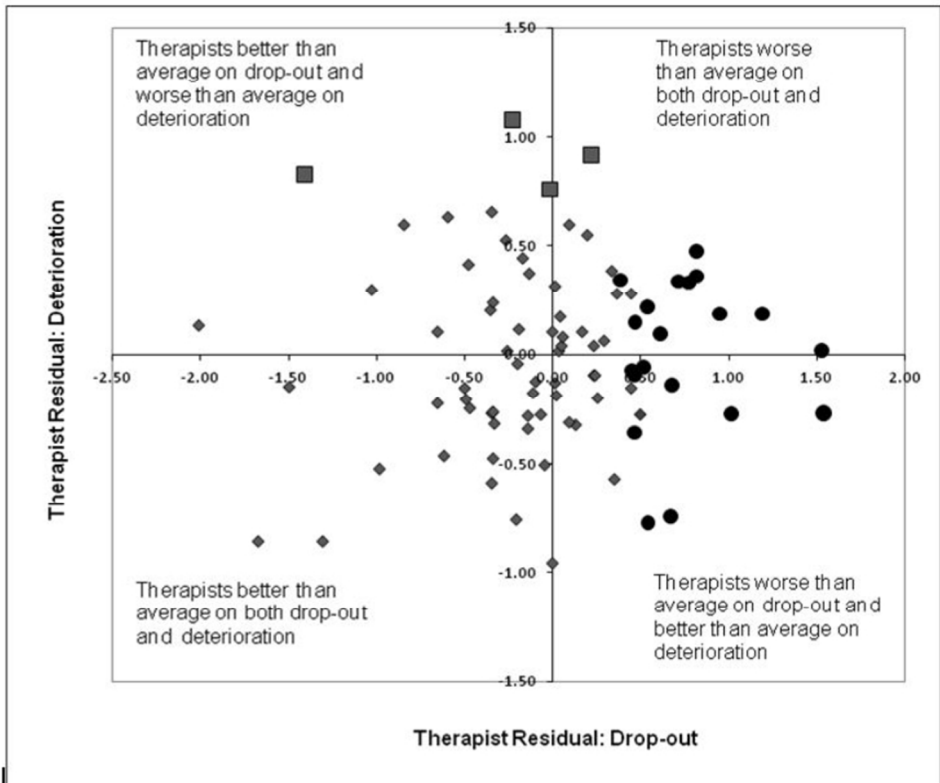


Figure 4: Scatterplot of therapist residuals for both dropout and deterioration
241x203mm (96 x 96 DPI)

view