

# Duration of psychological therapy: relation to recovery and improvement rates in UK routine practice<sup>†</sup>

William B. Stiles, Michael Barkham and Sue Wheeler

## Background

Previous studies have reported similar recovery and improvement rates regardless of treatment duration among patients receiving National Health Service (NHS) primary care mental health psychological therapy.

## Aims

To investigate whether this pattern would replicate and extend to other service sectors, including secondary care, university counselling, voluntary sector and workplace counselling.

## Method

We compared treatment duration with degree of improvement measured by the Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM) for 26 430 adult patients who scored above the clinical cut-off point at the start of treatment, attended 40 or fewer sessions and had planned endings.

## Results

Mean CORE-OM scores improved substantially (pre–post effect size 1.89); 60% of patients achieved reliable and clinically significant improvement (RCSI). Rates of RCSI and reliable improvement and mean pre- and post-treatment changes were similar at all tested treatment durations. Patients seen in different service sectors showed modest variations around this pattern.

## Conclusions

Results were consistent with the responsive regulation model, which suggests that in routine care participants tend to end therapy when gains reach a good-enough level.

## Declaration of interest

M.B. was a developer of the CORE measures.

## Copyright and usage

© The Royal College of Psychiatrists 2015.

In studies of routine mental health practice in UK National Health Service (NHS) primary care settings, patients who had scores above the clinical cut-off level on the Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM) at the outset of psychological therapy,<sup>1,2</sup> and had planned endings, had similar rates of recovery and improvement regardless of how long they remained in treatment across the 0–20 session range examined.<sup>3,4</sup> Similar results were reported among patients in an American university counselling centre.<sup>5</sup> These results seem unexpected in the context of conventional psychological therapy dose–effect curves, which suggest recovery and improvement rates should increase with the number of sessions delivered, albeit at a negatively accelerating rate.<sup>6–8</sup> These unexpected results might reflect participants' responsive regulation of treatment duration to fit patient requirements.<sup>3,4</sup> In the NHS primary care investigations the recovery rate was paradoxically slightly greater among patients who had fewer sessions than among those who had more sessions. Patients' pre-treatment and post-treatment CORE-OM scores were modestly correlated with the number of sessions they attended, suggesting that degree of distress was one of many factors influencing treatment duration. All scientific findings, but particularly such unexpected and paradoxical findings, must be reproduced across settings and closely assessed before they can be considered trustworthy.<sup>9,10</sup> Understanding the effects of psychological therapy not as a fixed function of dose but as responsively regulated could have important implications for policy regarding prescribed numbers of sessions. We assessed whether the dose–effect pattern previously observed in primary care mental health settings could be observed in other service sectors, including secondary care, university counselling centres, voluntary organisations and workplace counselling centres.

<sup>†</sup>See editorial, pp. 93–94, this issue.

## Method

We studied adult patients aged 16–95 years ( $n = 26\,430$ ) drawn from the CORE National Research Database 2011 (described below) who returned valid pre- and post-treatment assessment forms, began treatment in the clinical range, completed 40 or fewer sessions and were described by their therapists as having had a planned ending. Endings were considered as planned if the therapist and patient agreed to end therapy or a previously planned course of therapy was completed. The patients on average were 38.6 years old (s.d. = 12.9); 18 308 (69.3%) were women; 23 104 (87.4%) were White, 1092 (4.1%) were Asian, 913 (3.5%) were Black and 1321 (5.0%) listed other ethnicity or their ethnicity was not stated, not available or missing. Most patients were not given a formal diagnosis, but therapists indicated the severity of their patients' presenting problems using categories provided on the CORE assessment form. Multiple problems were indicated for most patients. Patients' problems rated as causing moderate or severe difficulty on one or more areas of day-to-day functioning included anxiety (in 56.3% of patients), depression (38.3%), interpersonal relationship problems (38.8%), low self-esteem (33.9%), bereavement or loss (23.0%), work or academic problems (23.0%), trauma and abuse (15.4%), physical problems (11.6%), problems associated with living on welfare (11.6%), addictions (5.4%) and personality problems (5.1%), as well as other problems cited for fewer than 5% of the patients.

The patients were treated over a 12-year period at 50 services in the UK, including six primary care services (8788 patients), eight secondary care services (1071 patients), two tertiary care services (68 patients), ten university counselling centres (4595 patients), fourteen voluntary sector services (5225 patients), eight workplace counselling centres (6459 patients) and two private

practices (224 patients). The patients were treated by 1450 therapists who treated from 1 to 377 patients each (mean 18.2, *s.d.* = 42.0). Therapist demographic and professional characteristics were not recorded. Assignment of patients to therapists and treatments was determined using the service's normal procedures. Therapists indicated their treatment approaches at the end of therapy using the CORE end of therapy form (described below). The most common approaches were integrative (41.2%), person-centred (36.4%), psychodynamic (22.8%), cognitive-behavioural (14.9%), structured/brief (14.6%), and supportive (14.0%). Therapists reported using more than one type of therapy with many (41.6%) of the patients.

## Measures

### Outcome measure

The CORE-OM is a self-report inventory comprising 34 items that address domains of subjective well-being, symptoms (anxiety, depression, physical problems, trauma), functioning (general functioning, close relationships, social relationships) and risk (risk to self, risk to others).<sup>1,2</sup> Half the items focus on low-intensity problems (e.g. 'I feel anxious/nervous') and half focus on high-intensity problems (e.g. 'I feel panic/terror'). Items are scored on a five-point scale (0–4), anchored 'not at all', 'only occasionally', 'sometimes', 'often' and 'all or most of the time'. (Six positively worded items are reverse-scored.) Forms are considered valid if no more than three items are omitted.<sup>2</sup> Clinical scores on the CORE-OM are computed as the mean of completed items multiplied by 10, so clinically meaningful differences are represented by whole numbers. Thus, scores can range from 0 to 40. The 34-item scale has a reported internal consistency of 0.94,<sup>1</sup> and test-retest correlations of 0.80 or above for intervals of up to 4 months in an out-patient sample.<sup>11</sup> The CORE-OM's recommended clinical cut-off score of 10 was selected to discriminate optimally between a clinical sample and a systematic general population sample.<sup>12</sup>

### Assessment and end of therapy forms

On the CORE assessment form,<sup>13</sup> completed at intake, therapists gave referral information, patient demographic characteristics and data on the nature, severity and duration of presenting problems using 14 categories: depression, anxiety, psychosis, personality problems, cognitive/learning difficulties, eating disorder, physical problems, addictions, trauma/abuse, bereavement, self-esteem, interpersonal problems, living/welfare and work/academic. On the end of therapy form therapists reported information about the completed treatment, including the number of sessions the patient attended, whether the ending was planned or unplanned, and which type or types of therapy were used.<sup>13</sup>

## Procedure

Details of data collection procedures were determined by each service's normal administrative procedures and were not recorded. Patients completed the CORE-OM during screening or assessment or immediately before the first therapy session. The post-treatment CORE-OM was administered at or after the last session. Therapists completed the therapist assessment form after an intake session and the end of therapy form when treatment had ended. Data collection complied with applicable data protection procedures for the use of routinely collected clinical data.

Anonymised data were entered electronically at the site and later transferred electronically to CORE Information Management Systems (CORE-IMS). Ethics approval for this study was covered by National Research Ethics Service application 05/Q1206/128 (Amendment 3).

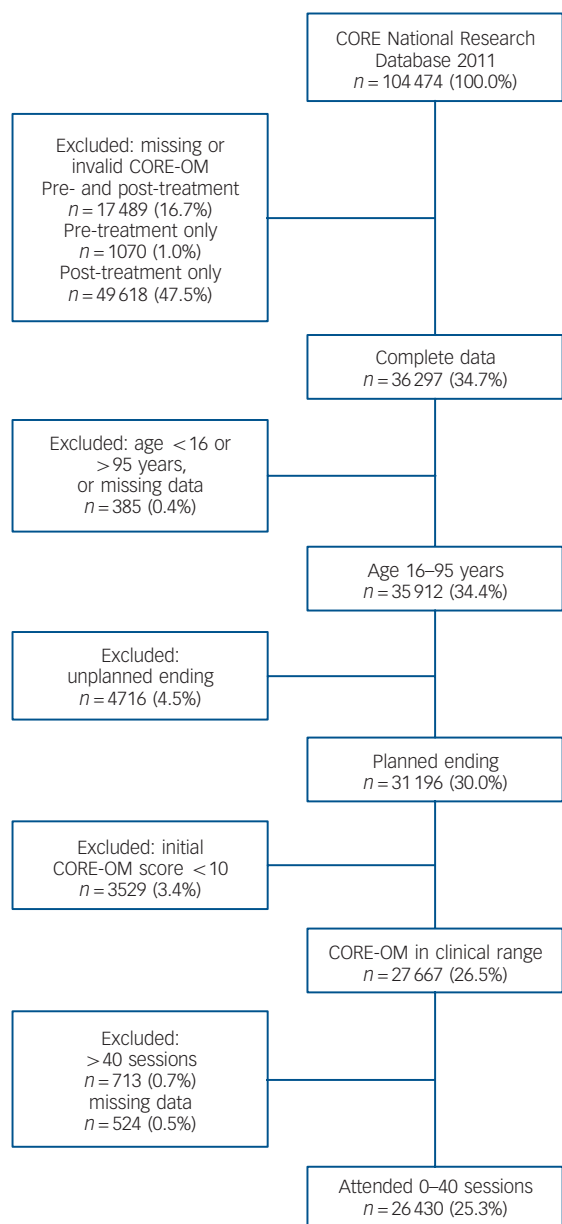
The 26 430 patients we studied were selected from the CORE National Research Database 2011 (CORE-NRD-2011). The CORE Information Management Systems invited all UK services using its software support services for more than 2 years (representing approximately 303 000 potential patients) to donate anonymised, routinely collected data for use by the research team. These services had been using the personal computer format of the CORE system.<sup>14</sup> Data were entered into the CORE-NRD-2011, which includes information on 104 474 patients (68.6% female; mean age 38.5 years, *s.d.* = 12.6) whose therapist returned a therapist assessment form during the period April 1999 to November 2011. For some patients data on more than one episode of therapy were donated; however, to ensure that each patient was included just once, only the first episode was included in the database (not necessarily the first episode the patient had ever had).

## Selection of patients

We selected all adult patients (age 16–95 years) from this database who returned valid pre- and post-therapy CORE-OM forms, were described by their therapists as having planned endings, had pre-treatment CORE-OM scores of 10 or greater, and completed 40 or fewer sessions. Criteria were decided *a priori*, paralleling the previous studies. We excluded patients in the database who did not return valid CORE-OM forms, including some who used short forms of the CORE rather than the CORE-OM: 17 489 did not return valid pre- or post-treatment forms, 1070 returned post-treatment but not pre-treatment forms and 49 618 returned pre-treatment but not post-treatment forms. The last (largest) category included patients who did not attend any sessions, patients who attended sessions but left without completing the final form and patients who had not ended their treatment by the closing date of data collection.

Of the 36 297 patients who returned valid pre- and post-treatment CORE-OM forms, 385 were excluded because there were missing data on age or age fell outside the 16–95 range; 4716 were excluded because they did not have a planned ending; and 3529 were excluded because they began treatment with a CORE rating below the clinical cut-off score of 10. Finally, we excluded 1237 patients who met previous criteria but received more than 40 sessions as reported on the end of therapy form, or whose therapist did not record the number of sessions attended. The number of patients receiving more than 40 sessions was too small to estimate recovery and improvement rates reliably. Many of the services routinely offered a fixed number of sessions, most often six. These limits were administered flexibly, however, and all services returned data from some patients seen for more than six sessions. This left our sample of 26 430, or 25.3% of the patients in the CORE-NRD-2011 (Fig. 1).

Some of the services asked to donate data to CORE-NRD-2011 had also been asked to donate data to previous data-sets, designated CORE-NRD-2002 and CORE-NRD-2005,<sup>3,4</sup> and there may have been some overlap with the present sample. Codes for patients, therapists and sites were assigned separately in each sample to preserve anonymity, so we could not check for overlap precisely. As a proxy estimate, we matched patients on five indices: age, pre-treatment CORE-OM score, post-treatment CORE-OM score, date of first session and date



**Fig. 1** Selection of patients from the Clinical Outcomes in Routine Evaluation (CORE) database. CORE-OM, CORE Outcome Measure.

of final session. In our study sample 472 of the 26 430 patients (1.8%) matched patients in CORE-NRD-2005 on these variables. This indicates an upper limit on the degree of overlap, insofar as all overlapping patients should match, and a few patients may have matched by chance. Of these, 469 were treated within the primary care sector, consistent with the restriction of the 2005 sample to primary care. The remaining 3 individuals, classified as workplace sector patients in CORE-NRD-2011, may have been coincidental matches, or the sector may have been misclassified in one of the data-sets. Most (80%) of the CORE-NRD-2011 patients began treatment after the end date for collection of the 2005 sample, so they could not have been included. There were many changes in which services were using CORE in 2005 and 2011, and services that donated data in 2005 may have chosen not to do so in 2011. A parallel check on overlap with the CORE-NRD-2002 sample showed no matches.

## Recovery

Reliable and clinically significant improvement

Reliable and clinically significant improvement (RCSI) is a common index of recovery in psychological therapy studies. Following Jacobson & Truax,<sup>15</sup> we held that patients had achieved RCSI if they entered treatment in a dysfunctional state and left treatment in a normal state, having changed to a degree that was probably not due to measurement error (that is, by an amount equal to or greater than the reliable change index described below). We used a CORE-OM score of 10 or above as the clinical cut-off level dividing the dysfunctional from the normal populations.<sup>12</sup> Recall that we excluded patients who began treatment with a score below the clinical cut-off; these patients were not in the clinical population to begin with and so could not move from the clinical to the normal population (as required for RCSI) regardless of how much they improved.

Reliable change index

The reliable change index is the pre–post treatment difference that, when divided by the standard error of the difference, is equal to 1.96. The standard error, and thus the index, depends on the standard deviation of the pre–post difference and on the reliability of the measure.<sup>15</sup> For comparability we used the same criteria as in the 2008 study by Stiles *et al.*<sup>4</sup> Using  $s.d._{diff} = 6.65$  (based on the 12 746 patients in CORE-NRD-2005 who had valid pre- and post-treatment CORE-OM scores) and the reported internal consistency reliability of 0.94, we calculated a reliable change index of 4.5. Values based on the 36 297 patients in CORE-NRD-2011 who returned valid pre- and post-treatment CORE-OM forms were similar:  $s.d._{diff} = 6.78$ , internal consistency averaged across pre- and post-treatment forms 0.94, reliable change index 4.49. To summarise, patients whose CORE-OM scores decreased by 4.5 points or more were considered to have achieved reliable improvement. Patients who achieved reliable improvement and whose score changed from at or above 10 to below 10 were considered to have achieved RCSI.

## Results

The 26 430 patients we studied attended an average of 8.3 sessions ( $s.d. = 6.3$ ; median and mode 6 sessions) and 15 858 of them achieved RCSI – that is, having begun treatment with a CORE-OM score of 10 or above, 60.0% of the patients left with a score below 10, having changed by at least 4.5 points. An additional 5258 patients (19.9%) showed reliable improvement only. Reliable deterioration (i.e. an increase of 4.5 or more points) was shown by 344 patients (1.3%). The remaining 4970 patients (18.8%) showed no reliable change. In this sample, with these parameters, the cut-off criterion (leaving treatment in a normal state) was a more powerful determinant of RCSI than the reliable change index (changing to a degree not attributable to chance). Whereas 21 116 (79.9%) of all patients showed reliable improvement (decrease greater than 4.5), only 16 585 (62.8%) ended treatment below the cut-off score of 10, just 727 (2.8%) more than achieved RCSI. The patients' mean CORE-OM clinical score was 18.99 ( $s.d. = 5.24$ ) at intake and 9.10 ( $s.d. = 6.28$ ) after treatment, a mean difference of 9.89 ( $s.d. = 6.48$ ),  $t(26,328) = 248.26$ ,  $P < 0.001$ , yielding a pre–post effect size (difference divided by pre-treatment  $s.d.$ ) of 1.89. Patients' pre-treatment scores were correlated ( $r = 0.38$ ,  $P < 0.001$ ,  $n = 26 430$ ) with their post-treatment scores.

## Dose–effect relations

Within the parameters of our study, patients who attended fewer sessions were somewhat more likely to have achieved RCSI by the end of treatment than were those who attended more sessions.

Table 1 shows the number and percentage of patients who achieved RCSI as a function of the number of sessions they attended (data for the full number of sessions are given in online Table DS1). All of the RCSI rates were between 45% and 70%, except for 37.5% among the relatively few patients who completed 39 sessions (online Table DS1). Rates of RCSI were negatively correlated with number of sessions attended ( $r = -0.58$ ,  $P < 0.001$ ; 41 session categories), replicating previous findings.<sup>3,4</sup> The proportion of patients achieving reliable improvement was also negatively correlated with number of sessions attended ( $r = -0.40$ ,  $P = 0.009$ ; 41 categories). Table 1 shows the number and percentage of patients who achieved reliable improvement as a function of number of sessions. Reliable improvement rates ranged from 70% to 85%, again except for a lower 57.5% among clients who attended 39 sessions (online Table DS1). These rates are higher than the RCSI rates because they included patients who achieved RCSI plus others who improved but did not end below the cut-off score. The 'no session' entry in Table 1 represents patients who returned to the site, perhaps after time on a waiting list, agreed with their therapist that formal treatment was no longer indicated, and completed a second CORE-OM. Although there were only a few such patients, their rate of improvement was consistent with the broader pattern, and we consider this to be a logical beginning of the continuum, reflecting treatment decisions by participants.

Mean pre-post treatment changes were similar regardless of the number of sessions attended. As shown in Table 2, the mean

pre-post change scores varied around the overall mean of 9.89, and the effect size varied around the overall effect size of 1.89. They were not significantly correlated with number of sessions across the 41 categories (0 to 40; online Table DS2). The patients who attended large numbers of sessions tended to have relatively higher CORE-OM scores both before and after treatment (Table 2, online Table DS2). The correlations of mean pre- and post-treatment CORE-OM scores with sessions attended across the 41 categories were substantial (pre-treatment  $r = 0.58$ ,  $P < 0.001$ ; post-treatment  $r = 0.66$ ,  $P < 0.001$ ). The correlations of individual patients' pre- and post-treatment CORE-OM scores with the number of sessions attended, although reliably positive, were far smaller, reflecting large within-category variation in CORE-OM scores (pre-treatment  $r = 0.08$ ,  $P < 0.001$ ; post-treatment  $r = 0.09$ ,  $P < 0.001$ ;  $n = 26\,430$ ).

### Service sector analysis

For the most part the foregoing pattern of results was observed within each of the service sectors we considered. Table 3 shows the RCSI rates and mean CORE-OM pre-post change in each of the sectors, with the whole-sample statistics included for comparison. Table 4 shows the mean number of sessions attended in each service sector within the sample we selected. We note that, descriptively, patients seen in secondary and tertiary care tended to begin with somewhat higher CORE-OM scores and to improve less than patients seen in other sectors, as indicated by relatively

**Table 1** Improvement rates as a function of number of sessions attended

Sessions attended <sup>a</sup>	Patients <i>n</i>	RCSI <sup>b</sup> <i>n</i> (%)	Reliable improvement <sup>c</sup> <i>n</i> (%)
0	28	19 (67.9)	22 (78.6)
1	243	152 (62.6)	185 (76.1)
2	1208	776 (64.2)	974 (80.6)
3	1926	1314 (68.2)	1609 (83.5)
4	2401	1603 (66.8)	2012 (83.8)
5	3071	1968 (64.1)	2538 (82.6)
6	5576	3225 (57.8)	4429 (79.4)
7	1904	1164 (61.1)	1563 (82.1)
8	1928	1128 (58.5)	1543 (80.0)
9	1107	620 (56.0)	877 (79.2)
10	1258	677 (53.8)	948 (75.4)
11	767	445 (58.0)	620 (80.8)
12	1405	798 (56.8)	1069 (76.1)
13	388	219 (56.4)	298 (76.8)
14	312	173 (55.4)	234 (75.0)
15	283	159 (56.2)	218 (77.0)
16	325	158 (48.6)	233 (71.7)
17	212	130 (61.3)	171 (80.7)
18	196	107 (54.6)	137 (69.9)
19	166	95 (57.2)	126 (75.9)
20	216	114 (52.8)	165 (76.4)
21	116	59 (50.9)	88 (75.9)
22	134	69 (51.5)	97 (72.4)
23	112	66 (58.9)	95 (84.8)
24	180	95 (52.8)	127 (70.6)
25	97	51 (52.6)	75 (77.3)
All patients	26 430	15 858 (60.0)	21 116 (79.9)

RCSI, reliable and clinically significant improvement.

a. Data for 26–40 sessions omitted here, but available in online Table DS1.

b. Defined as a post-treatment Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM) score below 10, which has changed by at least 4.5 points; includes only patients with planned endings whose initial CORE-OM score was at or above the clinical cut-off of 10.

c. Defined as a change of at least 4.5 points regardless of post-treatment score (and thus includes patients who achieved RCSI).

**Table 2** Clinical Outcomes in Routine Evaluation – Outcome Measure scores as a function of number of sessions attended

Sessions attended <sup>a</sup>	Patients <i>n</i>	CORE-OM score (mean)			
		Pre-treatment	Post-treatment	Pre–post difference	Pre–post effect size <sup>b</sup>
0	28	18.39	9.18	9.21	1.76
1	243	17.67	9.28	8.39	1.60
2	1208	18.04	8.38	9.66	1.84
3	1926	17.98	7.74	10.24	1.95
4	2401	18.44	8.10	10.34	1.97
5	3071	18.77	8.41	10.37	1.98
6	5576	19.05	9.29	9.76	1.86
7	1904	19.27	8.90	10.37	1.98
8	1928	19.24	9.23	10.00	1.91
9	1107	19.54	9.81	9.74	1.86
10	1258	19.24	10.19	9.05	1.73
11	767	18.99	9.29	9.71	1.85
12	1405	18.79	9.56	9.23	1.76
13	388	19.19	9.53	9.66	1.84
14	312	19.19	10.16	9.03	1.72
15	283	20.19	9.96	10.23	1.95
16	325	20.22	10.77	9.45	1.80
17	212	19.69	9.21	10.48	2.00
18	196	19.43	10.15	9.28	1.77
19	166	19.42	9.98	9.44	1.80
20	216	20.46	10.40	10.05	1.92
21	116	21.20	11.23	9.97	1.90
22	134	20.45	11.01	9.45	1.80
23	112	20.18	9.18	11.00	2.10
24	180	20.44	11.00	9.44	1.80
25	97	19.55	10.82	8.73	1.67
All patients	26 430	18.99	9.10	9.89	1.87

CORE-OM, Clinical Outcomes in Routine Evaluation – Outcome Measure.  
a. Data for 26–40 sessions omitted here, but available in online Table DS2.  
b. Pre–post treatment difference divided by the full sample pre-treatment standard deviation (5.24).

**Table 3** Changes in Clinical Outcomes in Routine Evaluation – Outcome Measure scores pre- and post-treatment categorised by mental health service sector

Service sector	Patients <i>n</i>	RCSI <i>n</i> (%)	Reliable improvement <i>n</i> (%)	Pre-treatment score Mean (s.d.)	Post-treatment score Mean (s.d.)	Pre–post difference <sup>a</sup> Mean (s.d.)	Effect size <sup>b</sup>
All clients	26 430	15 858 (60.0)	21 116 (79.9)	18.99 (5.24)	9.10 (6.28)	9.89 (6.48)	1.89
Primary	8788	5528 (62.9)	7258 (82.6)	19.36 (5.22)	8.73 (6.25)	10.63 (6.51)	2.03
Secondary	1071	386 (36.0)	707 (66.0)	21.47 (6.20)	13.75 (8.28)	7.72 (7.17)	1.47
Tertiary	68	18 (26.5)	35 (51.5)	20.17 (5.65)	14.41 (7.23)	5.76 (6.86)	1.10
University	4595	2740 (59.6)	3665 (79.8)	18.51 (5.12)	9.03 (5.57)	9.48 (6.13)	1.81
Voluntary	5225	2985 (57.1)	4032 (77.2)	18.76 (5.30)	9.46 (6.38)	9.29 (6.51)	1.77
Workplace	6459	4035 (62.5)	5221 (80.8)	18.58 (4.98)	8.56 (5.98)	10.02 (6.36)	1.91
Private	224	166 (74.1)	198 (88.4)	18.89 (5.00)	7.84 (5.26)	11.05 (5.82)	2.11

RCSI, reliable and clinically significant improvement.  
a. All pre–post differences were significant by paired *t*-test,  $P < 0.001$ .  
b. Mean difference divided by whole-sample pre-treatment standard deviation (5.24).

lower reliable change rates and effect sizes (see Table 3) despite having longer treatments (see Table 4). Private sector patients averaged somewhat more improvement; however, in view of the relatively small number of patients we are reluctant to interpret this. The durations of university and workplace treatments were similar to those in primary care; however, the durations of voluntary sector treatment were more similar to those in secondary care.

Figure 2 shows RCSI rates as a function of number of sessions completed for five of the sectors, with the whole-sample RCSI rates included for comparison. The whole-sample reference figure includes data for 0–40 sessions, as listed in online Table DS1. To ensure that the RCSI rates within sectors at each duration were based on a sufficient number of cases, we restricted the sector plots to primary care, secondary care, university counselling, voluntary sector and workplace counselling, and we included data

**Table 4** Number of sessions per patient categorised by mental health service sector

Sector	Patients <i>n</i>	Number of sessions Mean (s.d.)
Primary	8788	6.5 (3.4)
Secondary	1071	15.8 (9.7)
Tertiary	68	21.6 (10.0)
University	4595	6.4 (4.4)
Voluntary	5225	13.2 (8.5)
Workplace	6459	6.9 (4.4)
Private	224	7.3 (4.5)

only for patients who received 1 to 25 sessions. Like the whole-sample graph, all of the separate sector graphs show similar degrees of improvement for patients who received different numbers of sessions. Within this broad pattern, it is interesting to note that the modest negative correlation of RCSI rates with treatment duration reported previously,<sup>3,4</sup> and observed across sessions 0–40 in the full sample, also appeared numerically across the smaller range of sessions (1–25) in primary care ( $r = -0.33$ ,

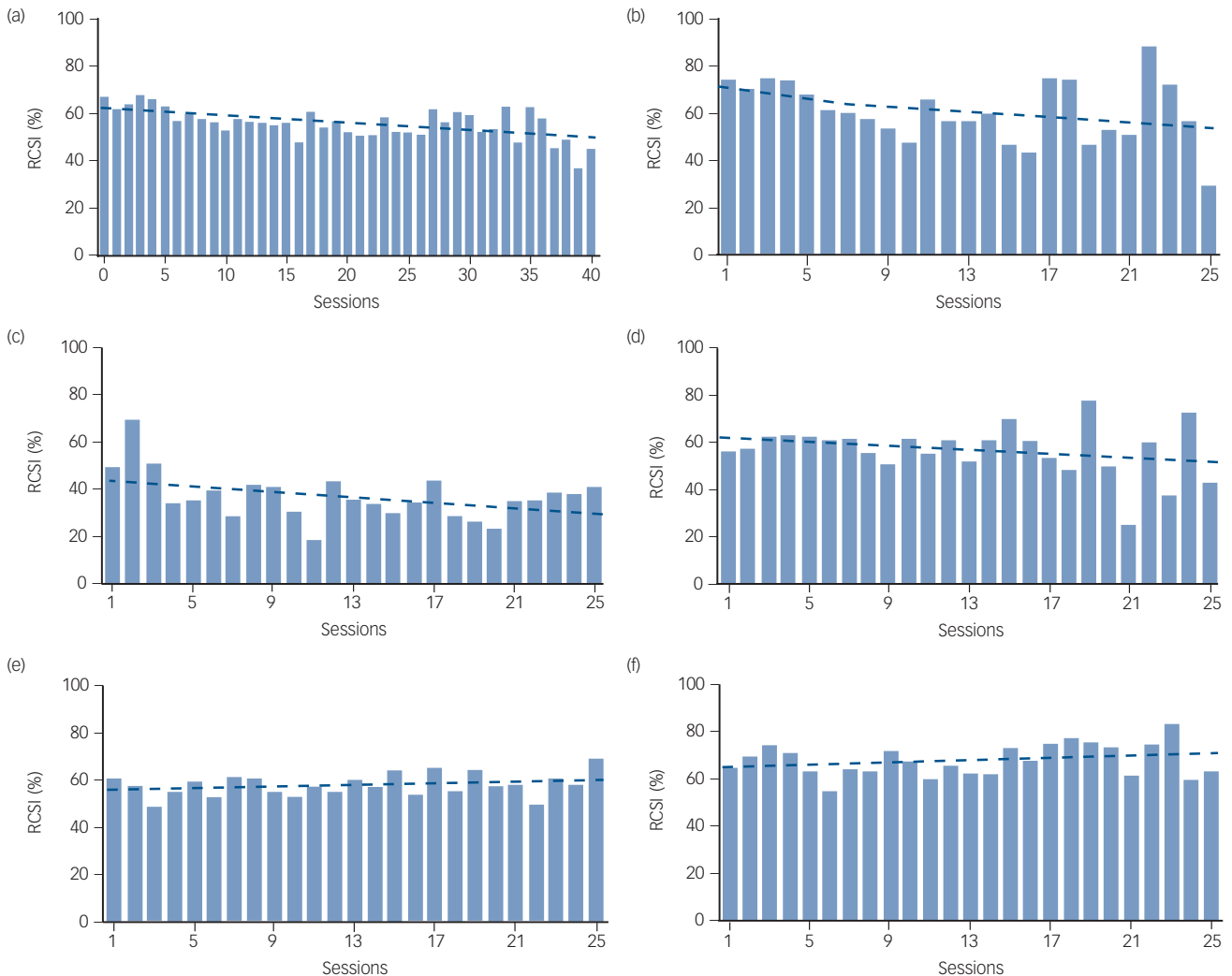
$P = 0.11$ ), secondary care ( $r = -0.40$ ,  $P = 0.05$ ) and university counselling ( $r = -0.28$ ,  $P = 0.18$ ), but not in the voluntary sector ( $r = 0.31$ ,  $P = 0.13$ ) or workplace counselling ( $r = 0.26$ ,  $P = 0.22$ ).

**Discussion**

As in the previous studies of therapy in primary care,<sup>3,4</sup> patients in this larger, more diverse sample averaged similar gains regardless of treatment duration. Confidence in the finding is strengthened by observing this pattern in service sectors other than primary care.

**Responsive regulation model**

Finding that patients seen for many sessions average no greater improvement than patients seen for few sessions may seem surprising if treatment duration is considered as a planned intervention or an independent variable in an experimental manipulation; but it may seem more plausible if patients and therapists are considered as monitoring improvement and adjusting treatment duration to fit emerging requirements, responsively ending treatment when improvement reaches a satisfactory level, given available resources and constraints. We call this ‘responsive



**Fig. 2** Reliable and Clinically Significant Improvement (RCSI) rate according to number of sessions attended in selected mental health service sectors: (a) whole sample; (b) primary care; (c) secondary care; (d) university counselling; (e) voluntary sector; (f) workplace counselling. With 7 exceptions each of the 165 categories illustrated represents 10 or more patients (in primary care 8 patients had 22 sessions, 7 had 23 sessions, 9 had 24 sessions and 7 had 25 sessions; in secondary care 4 patients had 1 session; in university counselling centre 8 patients had 23 sessions and 7 had 25 sessions).

regulation' of treatment duration.<sup>4,16</sup> The responsive regulation model presumes that patients have varied goals, potentials and expectations, change at different rates (for many reasons) and achieve sufficient gains at different times. In routine practice, participants (patients, therapists, family, administrators) respond to this variation by adjusting treatment duration so that treatments tend to end when patients have improved to a 'good-enough' level;<sup>3,17</sup> that is, improvement that is good-enough from the perspective of participants when balanced against costs and alternatives. Our results are consistent with this interpretation, although they do not prove it.

Despite diverging from the individual-level interpretation of dose–effect common in medicine, the responsive regulation model is consistent with the population-level interpretation of dose–effect curves, such as is used in agriculture – for example, to describe the percentage of weeds killed by a given dose of herbicide. Applied to psychological therapy, the population interpretation suggests that the easy-to-treat patients reach their good-enough level quickly and leave treatment, so a declining number of harder-to-treat patients remain in later sessions.

### Differences between service sectors

Gains were similar at different treatment durations within each of the mental healthcare sectors we studied. However, there were a few small but interesting differences between sectors. Patients in the secondary and tertiary care sectors averaged greater severity at intake and made fewer gains than patients in other sectors despite having longer treatments, presumably because they were referred for relatively chronic and complex problems judged likely to respond to treatment more slowly. Their lower level of gains suggests a lower good-enough level, which may reflect lower aspirations (e.g. coping with problems rather than full recovery), greater costs of care or the relative intractability of their problems to psychological treatment. Conversely, the RCSI rate of about 60% and other indices of gains were similar across the other sectors, suggesting that the good-enough levels were similar across these sectors.

The slight negative association of RCSI rates with treatment duration in primary and secondary care and university counselling replicated observations in the previous primary care samples.<sup>3,4</sup> However, this trend did not appear in either voluntary sector or workplace counselling. Previously, the negative correlation of treatment duration with RCSI rates in primary care has been interpreted as reflecting increasing costs of longer treatment;<sup>4</sup> that is, patients are satisfied with less as costs and difficulties mount. Costs include not only treatment fees but also time off work, child care, stigma of being in therapy and pressures from administrators to finish. We speculate that these cost increments are greater in the NHS and universities – large public-sector organisations with less freedom in their practice – than in the voluntary and workplace sectors.

The voluntary sector may be selectively treating cases that respond well to longer treatments: their RCSI rates were as strong as those of primary care, their treatment durations were as long as secondary care, and they showed no decline in outcomes at longer durations. The voluntary sector tends to serve people with relatively difficult problems with modest financial resources and good motivation for treatment.<sup>18</sup> We speculate that insofar as voluntary services impose fewer pressures to keep treatment short, people who are expected (by themselves or by others) to require longer treatments are directed towards them.

### Limitations

Naturalistic variation in treatment duration is not a natural experiment. Our naturalistic, practice-based design did not

randomly assign patients to receive different durations of treatment. Thus, our results do not suggest that any particular patient would have the same outcome regardless of treatment length. On the contrary, the responsive regulation model suggests that the optimum number of sessions varies with many aspects of patient, therapist and context. Results might be different in an experiment in which patients were randomly assigned to fixed numbers of sessions.<sup>19</sup> Theoretically, patients assigned to a duration below their optimum would be likely to finish below their good-enough level, whereas patients assigned to a duration above their optimum would receive more therapy than under naturalistic conditions and might accomplish more, perhaps finishing above their good-enough level.

### Non-completion of treatment

Our conclusions apply only to patients whose therapists said the ending was planned and who completed a post-treatment CORE-OM. Patients who left treatment before their planned ending seldom completed post-treatment measures, so we could not assess their gains. Patients may fail to appear for scheduled sessions because they have found other sources of help,<sup>20</sup> or feel that they have achieved their goals.<sup>21</sup> On average, however, patients who did not return to complete post-session measures seem likely to have made smaller gains than had patients who did return.<sup>22</sup> The large number of patients who did not complete post-treatment CORE-OM forms raises concerns about possible selective reporting. In a check on this possibility in the CORE-NRD-2005,<sup>23</sup> the proportion of patients who returned post-treatment forms varied hugely across the 343 therapists who saw 15 or more patients (this included 95% of the patients in that data-set) and across the 34 services they represented, but the rate of return was essentially uncorrelated with indices of improvement across therapists or across services. If therapists were selectively influencing patients with good outcomes to return post-treatment forms, improvement rates would be negatively correlated with reporting rates – that is, the more selective therapists would tend to have higher improvement rates. The observed lack of correlation suggests that therapists and services were not selectively reporting to any significant degree.

### Clinical implications

The responsive regulation model, which suggests that participants tend to regulate their use of therapy to achieve a satisfactory level of gains, shifts attention from decisions about treatment length to the question of what constitutes good-enough gains. Adjusting the cost and difficulty of access to and continuation of treatment, as well as influencing standards of what is good-enough by informing potential patients about what therapy can accomplish, might affect the average good-enough level. Our results suggest that the pattern of responsive regulation characterises diverse institutional settings. Allowing greater scope for responsive regulation might yield still greater efficiencies. For example, allowing patients to schedule their own appointments, rather than scheduling regular (e.g. weekly) appointments might dramatically reduce 'no-shows' without adverse effects on treatment effectiveness.<sup>24</sup>

**William B. Stiles**, PhD, Department of Psychology, Miami University, Oxford, Ohio, and Department of Psychology, Appalachian State University, Boone, North Carolina, USA; **Michael Barkham**, PhD, Centre for Psychological Services Research, Department of Psychology, University of Sheffield, UK; **Sue Wheeler**, PhD, Institute for Lifelong Learning, University of Leicester, UK

**Correspondence:** Dr William B. Stiles, PO Box 27, Glendale Springs, North Carolina 28629, USA. Email: stileswb@miamioh.edu

First received 28 Jan 2014, final revision 29 Oct 2014, accepted 3 Nov 2014

## Acknowledgements

We thank John Mellor-Clark and Alex Curtis Jenkins of Clinical Outcomes in Routine Evaluation Information Management Systems for collecting these data, making them available and documenting the procedures. We thank Sarah Bell for helping to put the data into a manageable format.

## References

- Barkham M, Gilbert N, Connell J, Marshall C, Twigg E. Suitability and utility of the CORE-OM and CORE-A for assessing severity of presenting problems in psychological therapy services based in primary and secondary care settings. *Br J Psychiatry* 2005; **186**: 239–46.
- Evans C, Connell J, Barkham M, Margison F, McGrath G, Mellor-Clark J, et al. Towards a standardised brief outcome measure: psychometric properties and utility of the CORE-OM. *Br J Psychiatry* 2002; **180**: 51–60.
- Barkham M, Connell J, Stiles WB, Miles JNV, Margison J, Evans C, et al. Dose-effect relations and responsive regulation of treatment duration: the good enough level. *J Consult Clin Psychol* 2006; **74**: 160–7.
- Stiles WB, Barkham M, Connell J, Mellor-Clark J. Responsive regulation of treatment duration in routine practice in United Kingdom primary care settings: replication in a larger sample. *J Consult Clin Psychol* 2008; **76**: 298–305.
- Baldwin SA, Berkeljon A, Atkins DC, Olsen JA, Nielsen SL. Rates of change in naturalistic psychotherapy: contrasting dose-effect and good-enough level models of change. *J Consult Clin Psychol* 2009; **77**: 203–11.
- Howard KI, Kopta SM, Krause MS, Orlinsky DE. The dose-effect relationship in psychotherapy. *Am Psychol* 1986; **41**: 159–64.
- Howard KI, Lueger RJ, Maling MS, Martinovich Z. A phase model of psychotherapy outcome: causal mediation of change. *J Consult Clin Psychol* 1993; **61**: 678–85.
- Lueger RJ, Howard KI, Martinovich Z, Lutz W, Anderson EE, Grissom GR. Assessing treatment progress of individual patients using Expected Treatment Response Models. *J Consult Clin Psychol* 2001; **69**: 150–8.
- Begley CG, Ellis LM. Raise standards for preclinical cancer research. *Nature* 2012; **483**: 531–3.
- Ioannidis JPA. Why most published research findings are false. *PLoS Medicine* 2005; **2**: 696–701.
- Barkham M, Mullin T, Leach C, Stiles WB, Lucock M. Stability of the CORE-OM and the BDI-I prior to therapy: evidence from routine practice. *Psychol Psychother* 2007; **80**: 269–78.
- Connell J, Barkham M, Stiles WB, Twigg E, Singleton N, Evans O, et al. Distribution of CORE-OM scores in a general population, clinical cut-off points and comparison with the CIS-R. *Br J Psychiatry* 2007; **190**: 69–74.
- Mellor-Clark J, Barkham M, Connell J, Evans C. Practice-based evidence and the need for a standardised evaluation system: informing the design of the CORE system. *Eur J Psychother Counsell Health* 1999; **3**: 357–74.
- Mellor-Clark J. National innovations in the evaluation of psychological therapy service provision. *Journal of Primary Care Mental Health* 2003; **7**: 82–5.
- Jacobson NS, Truax, P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 1991; **59**: 12–9.
- Stiles WB, Honos-Webb L, Surko M. Responsiveness in psychotherapy. *Clinical Psychology: Science and Practice* 1998; **5**: 439–58.
- Barkham M, Rees A, Stiles WB, Shapiro DA, Hardy GE, Reynolds S. Dose-effect relations in time-limited psychotherapy for depression. *J Consult Clin Psychol* 1996; **64**: 927–35.
- Moore S. Voluntary sector counselling: has inadequate research resulted in a misunderstood and underutilised resource? *Counselling and Psychotherapy Research* 2006; **6**: 221–6.
- Barkham M, Rees A, Stiles WB, Hardy GE, Shapiro DA. Dose-effect relations for psychotherapy of mild depression: a quasi-experimental comparison of effects of 2, 8, and 16 sessions. *Psychother Res* 2002; **12**: 463–74.
- Snape C, Perren S, Jones L, Rowland N. Counselling – why not? A qualitative study of people's accounts of not taking up counselling appointments. *Counselling and Psychotherapy Research* 2003; **3**: 239–45.
- Hunsley J, Aubry TD, Verstervelt CM, Vito D. Comparing therapist and client perspectives on reasons for psychotherapy termination. *Psychotherapy* 1999; **36**: 380–8.
- Stiles WB, Leach C, Barkham M, Lucock M, Iveson S, Shapiro DA, et al. Early sudden gains in psychotherapy under routine clinic conditions: practice-based evidence. *J Consult Clin Psychol* 2003; **71**: 14–21.
- Barkham M, Stiles WB, Connell J, Mellor-Clark J. Psychological treatment outcomes in routine NHS services: what do we mean by treatment effectiveness? *Psychol Psychother* 2012; **85**: 1–16.
- Carey TA, Tai SJ, Stiles WB. Effective and efficient: using patient-led appointment scheduling in routine mental health practice in remote Australia. *Prof Psychol Res Pr* 2013; **44**: 405–14.



100  
words

## Romantic suicide

Alexandra Pitman

The concept of 'romantic suicide' emerged in the 1770s following the suicide of the teenage English poet Thomas Chatterton, memorialised in Goethe's 1774 novel in the character of young Werther. Eighteenth-century authors embellished these men as romantic outcasts, triumphing over death through fearless individualism to achieve immortality in heaven. Such myths still persist today, exemplified in journalists' responses to the suicides of artists such as Kitaj, Kirchner, Rothko and Van Gogh. Glorifying their deaths by wreathing them in martyrdom is a dangerous practice. Awareness of the damage wreaked by propagation of these myths is the starting point for challenging them.

The British Journal of Psychiatry (2015)  
207, 122. doi: 10.1192/bjp.bp.114.162370