Behavioural and Cognitive Psychotherapy, 2014, 42, 747–759
First published online 2 January 2014 doi:10.1017/S1352465813000544

Service Use, Drop-Out Rate and Clinical Outcomes: A Comparison Between High and Low Intensity Treatments in an IAPT Service

Stella W. Y. Chan

University of East Anglia, Norwich and University of Edinburgh, UK

Malcolm Adams

University of East Anglia, Norwich, UK

Background: The IAPT services provide high and low intensity psychological treatments for adults suffering from depression and anxiety disorders using a stepped care model. The latest national evaluation study reported an average recovery rate of 42%. However, this figure varied widely between services, with better outcomes associated with higher "step-up" rates between low and high intensity treatments. **Aims:** This study aimed to compare the two intensity groups in an IAPT service in Suffolk. **Method:** This study adopted a between groups design. A sample of 100 service users was randomly selected from the data collected from an IAPT service in Suffolk between May 2008 and February 2011. The treatment outcomes, drop-out rate, and other characteristics were compared between those who received high and low intensity treatments. **Results:** The high intensity group received, on average, more sessions and contact time. They received more CBT sessions and less guided self-help. There were no group differences in terms of the drop-out and appointment cancellation rates. Analyses on clinical outcomes suggested no group difference but demonstrated an overall recovery rate of 52.6% and significant reduction in both depression and anxiety symptoms. **Conclusions:** Despite methodological limitations, this study concludes that the service as a whole achieved above-average clinical outcomes. Further research building upon the current study in unpacking the relative strengths and weaknesses for the high and low intensity treatments would be beneficial for service delivery.

Keywords: IAPT, treatment outcome, drop-out, depression, anxiety, CBT

Introduction

Background of the IAPT Programme

The Improving Access to Psychological Therapies (IAPT) programme was an initiative launched by the UK Government in 2008 aiming at improving access to evidence-based

Reprint requests to Stella W. Y. Chan, Clinical Psychology, School of Health in Social Science, University of Edinburgh, Edinburgh EH8 9AG, UK. E-mail: stella.chan@ed.ac.uk

psychological treatments for people who are suffering from depression and anxiety disorders. The driving force behind this large-scale scheme was the publication of the *Depression Report* by Lord Layard and his colleagues (Layard et al., 2006), in which the authors argued, from both economic and clinical perspectives, the urgent need for improving the provision of psychological treatments. Specifically, the report highlighted the alarming figures concerning the scale of mental illnesses, their disabling effects, and notably the fact that 75% of those diagnosed are left untreated in part due to the long waiting time and the patchy nature of the psychological therapy provision. Following the argument that the funding for providing more psychological therapies would eventually pay for itself by reducing other mental illness related public costs (e.g. welfare benefits, medical costs, employment rates and hence revenue from taxes), the report saw the set-up of two demonstration sites in Doncaster and Newham.

Following the success in these two demonstration sites (Clark et al., 2009), the nationwide roll-out of IAPT services was launched in full; by March 2011, 142 of the 151 Primary Care Trusts in England had a service from this programme in at least part of their area, providing access to over 50% of the adult population. To support a programme of this scale, a range of training programmes have been set up and so far 3660 cognitive behavioural therapy (CBT) workers have already been trained (www.iapt.nhs.uk). Although initially set up to target adults of working age, the programme has been open to adults of all ages since 2010.

High vs. low intensity interventions

One of the primary objectives of IAPT is to facilitate consistent implementation of the National Institute for Health and Clinical Excellence (NICE) Guidelines in the treatment for depression and anxiety disorders (www.iapt.nhs.uk). Thus, the IAPT services provide exclusively evidence-based interventions such as CBT (Clark and Ehlers, 1993; Shapiro et al., 1994) delivered through a stepped care model. As noted above, access to psychological interventions has been limited, in part due to the restricted resources available; a stepped care model was proposed as a way to improve access through more efficient use of therapy resources (e.g. Scogin, Hanson and Welsh, 2003). This model was developed based on two principles: first, the model aims to provide interventions at the right time and as adequately as possible, i.e. "not earlier or more intensely than necessary" and "not later or less intensely than needed"; second, the model is "self-correcting", meaning that therapeutic progress is systematically monitored such that service users will be "stepped-up" to more intense interventions if they do not respond adequately to the less intense treatments (Seekles, van Straten, Beekman, Marwijk and Cuijpers, 2011). The rationale was that, in this way, limited resources could be used more efficiently by having the more costly interventions reserved only for those who need them (Bower and Gilbody, 2005).

Advocating a stepped care model, the IAPT services range between Step 1 (assessment and watchful waiting), Step 2 (low intensity interventions) and Step 3 (high intensity interventions). Low intensity treatments, usually provided by psychological well-being practitioners, include computerized CBT (cCBT), guided and/or pure self-help, behavioural activation, exercise, and psycho-education groups. High intensity usually refers to face-to-face CBT delivered by suitably trained professionals, although other therapeutic approaches such as Interpersonal Psychotherapy, Couples Therapy and Counselling are also available. These treatment choices are in line with the NICE Guidelines (NICE, 2004a, b), which also suggest that service users with mild to moderate clinical presentations should initially be

offered the "least burdensome" treatment they require, which could be stepped-up to higher intensity approaches if necessary. However, specific types of anxiety disorders, such as social phobia and post-traumatic stress disorder, and more severe depression are recommended to commence high intensity treatment from the start, which could be stepped-down to lower intensity approaches when appropriate.

Outcome evaluation and recovery rates

With its strong emphasis on quality assurance, IAPT services routinely record their activity for monitoring and evaluating purposes. Information such as demographic data of service users, sources of referrals, types of treatments, and service use are thus well documented. Specifically, the progress of each client is monitored through repeated measurement of depression and anxiety symptoms in each session using standard psychological questionnaires.

Based on the data collected from the 32 (out of 35) services established in the first year (October, 2008 – September, 2009), 64% of the service users were reported to show "reliable" improvement in their symptom scores (i.e. a reduction in scores that was beyond the expected error of the measurement). Notably, 42% amongst those who had attended at least two sessions (including assessment) were reported to have reached the pre-determined criteria for recovery (i.e. had both their depression and anxiety symptoms scored below clinical cut-off points) (North East Public Health Observatory (NEPHO), 2010).

While the average recovery rate was encouraging, a considerable variability (between 27% and 58%) was found between sites (NEPHO, 2010). Further analyses suggested that better clinical outcomes were reported amongst services that offered a larger number of treatment sessions, had a higher proportion of experienced therapists, and showed better compliance with NICE treatment recommendations (Gyani, Shafran, Layard and Clark, 2011). Notably, this report also suggested that services that made better use of stepped care produced better outcomes, and further estimated that an improvement on the "step-up" rate could potentially increase the recovery rate to as high as 54%.

Given this large variability between services, it is important to explore the data from individual services separately. Findings from these analyses would be helpful in improving future service development. Specifically, the finding that recovery rate was associated with step-up rate warrants further research into the differences between high and low intensity treatments provided within the same service.

Aims of the study

In light of the above, this study examined the data collected from an IAPT service in Suffolk between May 2008 and February 2011, with a specific aim to compare those who received high intensity versus low intensity treatments in terms of the following:

1. The type and duration of treatments they received;
2. The rate of drop-out and appointment cancellation;
3. The outcome of treatment.

https://doi.org/10.1017/S1352465813000544 Published online by Cambridge University Press

Method

Overview

This study adopted a between groups design. A large dataset was obtained from the IAPT service in Suffolk, from which a sample of 100 service users was randomly selected, with half of them receiving high intensity and half receiving low intensity interventions. Their characteristics were compared. See below for details.

Service context

The IAPT service in Suffolk provides support to individuals with stress-related difficulties as well as those with mild to moderate symptoms of depression or anxiety disorders. Service users could be self-referred or referred by professionals (e.g. GP). Upon receipt of a referral, a member of the team (usually a psychological wellbeing practitioner) will conduct an initial assessment with the client to determine suitability of the service to the client's reported difficulties. In line with the stepped care model, the majority of clients will be offered low intensity interventions, such as guided self-help, as the starting point. Clients not responding to these interventions or judged clinically to require more intensive therapeutic work will be stepped-up to high intensity interventions, which consist of mainly face-to-face CBT. In addition, clients with social phobia or PTSD will be stepped-up to high intensity treatments directly from assessment. In terms of staff training level, low intensity treatments are delivered by psychological wellbeing practitioners who have had one year training in delivering guided self-help. High intensity treatments are delivered by accredited therapists.

The protocols for non-attenders and prevention of drop-out are as follows: if a client did not attend (DNA) the first assessment session, the client would not be offered a further appointment. If a client DNA a treatment session, a letter would be sent to ask the client to get in touch with the service to arrange another appointment. If the client did not get in touch within 2 weeks, it would be assumed that the client no longer required the service. If a client DNA two treatment sessions, no further appointment would be offered. To enhance engagement and reduce drop-out, clinical staff routinely talk to clients about their readiness to change and level of motivation, especially at the start of the therapy.

Sample

The original dataset contained information about 15,082 service users, including their demographic data (e.g. age, gender, ethnicity), referral information (e.g. referral source, date, and reason), and the assessment and treatments they received at IAPT. Each client contact was recorded as a separate entry, giving rise to an alarming figure of 88,072 entries in total. It was deemed infeasible, within the scope of this study, to examine the entire dataset. The following analyses therefore included data of only 100 service users randomly selected from the original dataset. This sample size was determined based on general statistics principles (Howell, 2002); a formal power calculation was impossible due to the unavailability of relevant data in the existing literature.

Specifically, 50 were selected from each of the two groups (i.e. high intensity group vs. low intensity group). The intensity level was not explicitly stated in the original dataset. However, following a discussion with the service, it was decided that those who were seen by more

https://doi.org/10.1017/S1352465813000544 Published online by Cambridge University Press

Table 1. Demographic and referral information of the high and low intensity groups

Variable	Category	High intensity %	Low intensity %
Gender	Male	30	30
	Female	68	68
	Not stated	2	2
Ethnicity	British	62	78
	Other White	2	4
	Not stated*	36	18
Referral source	GP	78	88
	Self	16	6
	Other	6	6
Reason for referral (primary diagnosis)	Depression	16	30
	GAD	24	10
	Phobia	2	6
	PTSD	6	0
	OCD	8	2
	Mixed depression and anxiety	32	38
	Other/Not stated	12	14

Notes: *denotes statistically significant group difference $p < .05$
GAD = Generalized Anxiety Disorder; PTSD = Post-Traumatic Stress Disorder; OCD = Obsessive Compulsive Disorder

senior therapists (band 6 - 8) would be classified into the high intensity group whereas those who were seen by more junior staff members (band 4 - 5) would be classified into the low intensity group.

It should be noted that, with the randomization procedure used in this study, clients who received more clinical contacts had a higher probability of being selected as they had more data entries. While reducing the randomness of the selection, this method resulted in a sample that was representative of the actual workload and resource distribution of the service and thus was deemed to be a desirable option. The potential limitation of this randomization procedure will be further considered below in Discussion.

The demographic and referral details of the two groups are summarized in Table 1. Overall, about two-thirds of the sample was female and the majority stated their ethnicity as British. More than three-quarters of the clients were referred by their GPs, with primary diagnoses spread across depression, different types of anxiety disorders, and mixed depression and anxiety disorders. Significantly, more clients from the low intensity group had their ethnicity recorded as "not stated" ($z = 2.03$, $p = .04$). Apart from this, the two groups did not differ in any other characteristics (all z values < 1.96).

Age was recorded as ranges (e.g. 20–29), and in this study the actual age was estimated by taking the midpoint of the range (e.g. 25). There was no group difference in the average age (low intensity 44.1 ± 11.5 vs. high intensity 43.5 ± 15.0, $t\,(98) = .20$, $p = .84$).

Measures

As part of the routine practice at IAPT services, the severity of depressive and anxiety symptoms of each client were assessed in every session using the Depression Module of

https://doi.org/10.1017/S1352465813000544 Published online by Cambridge University Press

the Patient Health Questionnaire (PHQ-9; Kroenke, Spitzer and Williams, 2001) and the Generalized Anxiety Disorder Assessment (GAD-7; Spitzer, Kroenke, Williams, and Lowe, 2006) respectively.

Both are brief, self-reported measures. The PHQ-9 contains 9 items that correspond to the diagnostic criteria for depression defined in the *Diagnostic and Statistical Manual for Mental Disorders (4th edn.)* (DSM-IV; American Psychiatric Association, 2000), and requires clients to rate how often they experience each of the symptoms in the past 2 weeks on a scale of 0 (not at all) to 3 (nearly every day). The GAD-7 was originally designed to assess severity of Generalized Anxiety Disorders (GAD) but has also been commonly used to assess other types of anxiety disorders. The GAD-7 contains seven items, most of which were drawn from the DSM-IV criteria for GAD, and requires clients to give ratings similar to that in the PHQ-9.

A good internal consistency (Cronbach's $\alpha = .92$) and test-retest reliability ($r = .83$) were reported for GAD-7, which was also shown to have a strong correlation ($r \geq .72$) with other anxiety measures such as the Beck Anxiety Inventory, suggesting a good construct validity (Spitzer et al., 2006). Similarly, the PHQ-9 was reported to have an internal consistency of .86, a test-retest reliability of .84, a sensitivity of 88% and a specificity of 88% for major depression (Kroenke et al., 2001).

Data analyses

Analyses were conducted using SPSS. Parametric assumptions were verified prior to performing the following statistical tests. Independent samples *t* tests were used to examine group differences in terms of the type and duration of treatments, as well as the number of appointments cancelled by therapists or clients. The frequency of drop-out was compared using a *Z* test. Treatment outcome was measured in two ways. First, the change in PHQ-9 and GAD-7 scores from initial assessment to the final therapy session was investigated using two separate Repeated Measures ANOVAs, with intensity as the between-subjects variable (two levels: high vs. low) and time as the within-subjects variable (two levels: pre- vs. posttreatment). Any significant interactions would be clarified by individual *t* tests. Effect sizes of these changes were also calculated. Second, a *Z* test was used to compare the two groups in terms of the proportion of service users who reached the predetermined criteria for recovery by the end of the therapy. Service users were classified as "recovered" if both of their depression and anxiety symptoms were scored below the clinical cut-off points for moderate presentations (i.e. PHQ-9 ≤ 9 and GAD-7 ≤ 7), and that only those who attended at least two sessions including the assessment were included in this analysis. These outcome measures were chosen to be consistent with previous evaluation studies (Clark et al., 2009; NEPHO, 2010). The final therapy scores were not recorded for nine service users, and on these occasions, their last recorded scores were used. In the cases in which only one session took place, the initial and final scores were the same. Statistical significance was defined as $p < .05$ (two tailed).

Ethical considerations

This study was approved by the University of East Anglia (Faculty of Health) Research Ethics Committee. All data were anonymized and kept confidential according to the requirements of the Data Protection Act. This study made use of the data collected by the service as part of

Table 2. Number of sessions, contact time, and appointment cancellation of the high and low intensity groups

Variable	High intensity	Low intensity	*t*	*p*
Total no. of sessions	15.98 (8.95)	6.61 (3.99)	6.69	< 0.001*
Total contact time (hours)	12.42 (7.74)	2.59 (2.47)	8.47	< 0.001*
No. of sessions cancelled by therapist	0.55 (0.78)	0.33 (0.63)	1.57	.119
No. of sessions cancelled by client	2.45 (2.37)	1.76 (1.64)	1.68	.096

Notes: Values represent group means and standard deviations (in bracket).
*denotes statistically significant group difference $p < .05$

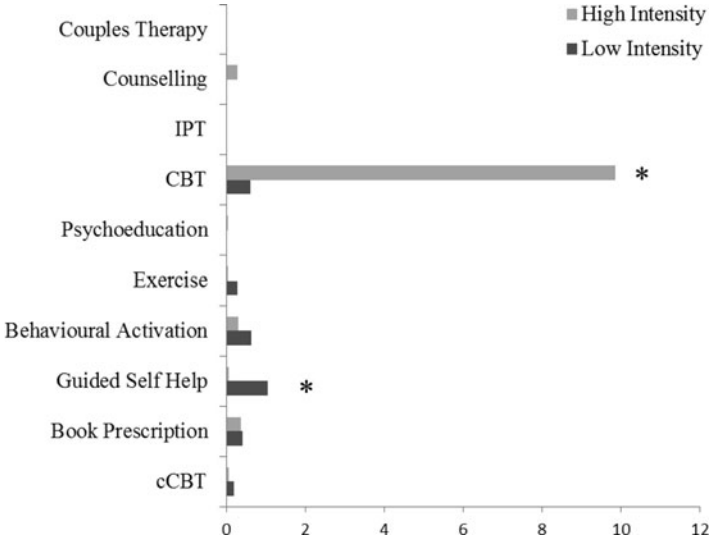

Figure 1. Average number of sessions for each type of treatment
Notes: Bars represent group means of the high and low intensity groups; *denotes statistically significant group difference $p < .05$

its routine practice and therefore it did not inflict an additional burden on service users or the staff team. Results of the study were disseminated to the service involved, with the aim to help to improve their future service provision.

Results

Type and duration of treatments

As expected, the high intensity group received more sessions and total contact time than the low intensity group (all p values < .001; see Table 2). The average number of sessions for each type of treatment received by the two groups is shown in Figure 1. Statistical analyses show that the high intensity group received more CBT sessions but less guided self-help than the low intensity group ($p < .001$). The two groups did not differ in the frequency of other

https://doi.org/10.1017/S1352465813000544 Published online by Cambridge University Press

Table 3. Percentage of clients attending at least one session of each type of treatment

Type of treatment	High intensity %	Low intensity %	z
cCBT	2	6	1.02
Book	24	24	0.00
Guided self-help	6	38	3.86*
Behavioural activation	14	26	1.50
Exercise	2	12	1.96*
Psycho-education	4	0	1.43
CBT	86	16	7.00*
IPT	0	0	N/A
Counselling	2	2	0.00
Couples therapy	0	0	N/A

Note: *denotes statistically significant group difference $p < .05$.

Table 4. Number of clients classified by reason for end of therapy ($n = 50$ per group)

Reason for end of therapy	High intensity	Low intensity	z
Completed treatment	22	16	1.15
Declined treatment	5	9	1.15
Dropped out	3	9	1.85
Not suitable for service	3	1	1.02
Not known	17	15	0.43

types of treatments (all p values $> .10$). In addition, Table 3 presents the percentage of clients in each group who received at least one session of each type of treatment. Consistent with the above, more clients from the high intensity group received CBT ($p < .001$) but fewer clients from this group received guided self-help ($p < .001$) or exercise ($p = .05$).

Rates of drop-out and appointment cancellation

There was no group difference in terms of the number of sessions cancelled by therapist or by client (both p values $\geq .10$; see Table 2). The reason for end of therapy was recorded in the dataset. The number of clients in each category for the two groups is summarized in Table 4. None of the group comparisons reached statistical significance.

Treatment outcomes

Depression and anxiety scores were missing from nine service users (two from the high intensity group and seven from the low intensity group), and therefore they were not included in the following analyses. In terms of the change in depressive symptoms as measured by PHQ-9 scores, there was a significant effect of time ($F(1, 89) = 52.39$, $p < .001$) suggesting

https://doi.org/10.1017/S1352465813000544 Published online by Cambridge University Press

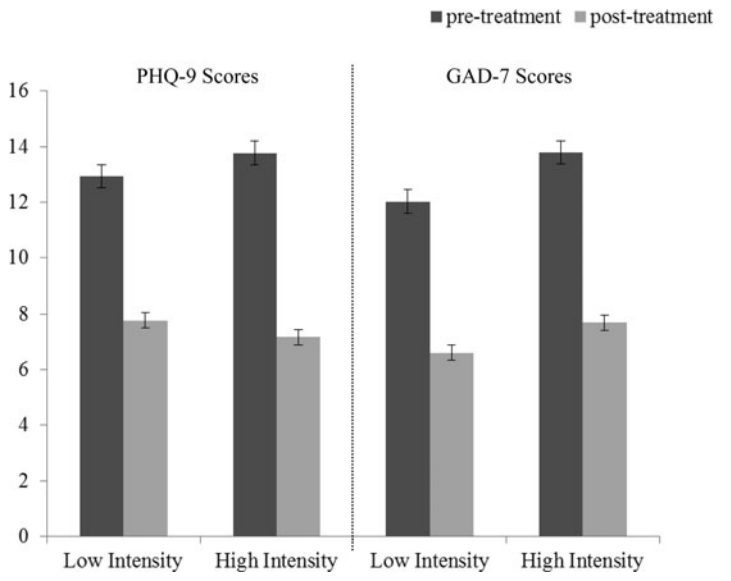

Figure 2. Changes in PHQ-9 (left) and GAD-7 (right) scores from before to after treatment
Note: Bars represent group means ± SEM of the low and high intensity groups before and after treatment

that this score reduced from pre- to posttreatment over the entire sample. However, there was no significant effect of group ($F(1, 89) = .01$, $p = .92$) or time x group interaction ($F(1, 89) = .79$, $p = .38$). Similar results were found for the GAD-7 scores. While the overall sample showed a decrease in GAD-7 scores as suggested by a significant effect of time ($F(1, 89) = 76.95$, $p < .001$), there was no significant effect of group ($F(1, 89) = 2.21$, $p = .14$) or interaction ($F(1, 89) = .27$, $p = .60$). See Figure 2.

The effect sizes for the above changes in the depression and anxiety scores range between 0.66 and 0.99, corresponding to medium to large effect sizes (Cohen, 1988). Specifically, the effect sizes for the change in PHQ-9 scores were 0.77 (whole sample), 0.86 (high intensity group) and 0.66 (low intensity group). The corresponding effect sizes for the change in GAD-7 scores were 0.93, 0.99 and 0.85.

In terms of recovery rates, five service users (three from the low intensity group and two from the high intensity group) attended fewer than two sessions and therefore were excluded from the analysis of this outcome variable. Overall, 52.6% of the entire sample reached the criteria for recovery. The high and low intensity groups achieved a recovery rate of 50.0% and 55.3% respectively, with no significant group difference ($z = 0.52$, $p = .60$).

Discussion

This study compared the service use, drop-out rate, and clinical outcomes of service users who received high versus low intensity treatments delivered in routine clinical practice at an IAPT service. Results showed that the high intensity group received, on average, more sessions and

https://doi.org/10.1017/S1352465813000544 Published online by Cambridge University Press

total contact time. In particular, they received more CBT sessions and less guided self-help than did the low intensity group. These group differences could not be accounted for by drop-out or appointment cancellation rates, which were similar between the two groups. These group differences are consistent with the expectation that the high intensity interventions take up more resources from the service both in terms of time and human resources, especially from the personnel with higher levels of training, skills and experience.

Analyses on clinical outcomes revealed encouraging results in that both depression and anxiety symptoms reduced significantly across the whole sample, with medium to large effect sizes, although these figures should be treated with caution given that this study was not a randomized controlled trial. Nevertheless, the overall recovery rate of 52.6% is within the higher range of the national data (27% - 58%) and, notably, considerably higher than the national average of 42% (NEPHO, 2010). It is also positive to note that these clinical outcomes are consistent with the established effectiveness of CBT in previous studies (Clark and Ehlers, 1993), suggesting that the effectiveness of the treatments has not been compromised by practical constraints in routine practice at IAPT services.

However, it is slightly surprising that the two groups did not differ in their baseline depression and anxiety scores, given that the high intensity interventions were intended for clients with more severe clinical presentations. This could be because either the two groups differed in ways that were not captured by the questionnaires (PHQ-9 and GAD-7) or that they indeed did not differ.

Although both groups improved after treatment, there was no evidence for a difference between the high and low intensity groups. There are at least three possible explanations. First, as suggested above, the groups might differ in some ways uncharacterized by the present measurements. In particular, those in the high intensity group could experience clinical presentations that are harder to treat and/or resistant to change (e.g. recurrent depression). Second, service users who did not show progress from low intensity interventions were likely to have stepped-up to high intensity interventions. Thus, those who were less likely to reach "recovery" were included in the high intensity group, dragging down the recovery rate in this group. This hypothesis could not be tested directly due to the absence of the stepping-up data; further studies should ensure that indicators for stepping-up are properly identified and recorded for analysis. In addition, data regarding the use of medication have not been collected, rendering it impossible to examine its potential implications on the current results. Finally, the two groups may indeed be similar both before and after treatments. If this were the case, it would imply that the high intensity does not offer additional benefit, and further research should seek to test this hypothesis directly.

Critique of the methodology

The above results should be considered in the context of the methodological strengths and weaknesses. First, the findings should be interpreted in the context of the relatively small sample size; the reliability of the results could be enhanced by including a larger proportion of the original dataset. However, as discussed above, this was deemed infeasible due to the way the original dataset was organized.

To ensure that the selected service users were representative of the population of interest, randomization was used. As noted in the Method section, with the randomization procedure used in this study, clients who received more clinical contacts had a higher probability of

https://doi.org/10.1017/S1352465813000544 Published online by Cambridge University Press

being selected as they had more data entries. This method resulted in a sample that was representative of the actual workload and resource distribution of the service and thus was deemed to be a desirable option at the time of the study. However, on reflection, this bias should have been avoided given that the number of sessions and drop-out rates are key variables of interest.

IAPT has a strong emphasis on quality assurance, and clinical activities of these services are routinely recorded. Indeed, the current dataset provided information with a scale of richness unrivalled by other (more conventional) health services. The repeated measurement of depression and anxiety symptoms using validated psychological questionnaires was particularly helpful for evaluating treatment outcomes. However, there were occasions in which data appeared to be missing. In particular, there was insufficient information in this dataset to explore the rate of stepping-up (or stepping-down) in this service and its impact. Indeed, it is of particular importance that further research is done to clarify the effects of stepped care model. Previous research studies that investigated the benefits of a stepped care model for depression and anxiety have yielded mixed results. Although there has been research suggesting that services that made better use of the stepped care system achieved higher recovery rates (Gyani et al., 2011), two randomized-controlled trials found that stepped care model is not more beneficial than care as usual for clients with depression or anxiety disorders (Seekles et al., 2011; van Straten, Tiemens, Hakkaart, Nolen and Donker, 2006). A literature review also highlighted a number of methodological limitations in studies that attempted to evaluate the benefits of stepped care model, in particular the lack of research that examined the stepped care process as a whole (Bower and Gilbody, 2005). Taken together, the evidence base for stepped care model appears to be far from robust. Therefore, it is important to note that the clinical outcomes achieved by the IAPT service as shown in this report may not necessarily be accounted for by the stepped care model per se. Indeed, the benefit of IAPT could be that the stepped care acts as part of a co-ordinated system that helps reducing treatment delays.

The present study specifically investigated the drop-out rates and appointment cancellation rates, which were of particular interest as they give an indication of the waste of resources. However, the drop-out rate was determined by one variable in the original dataset known as "reason for end (of therapy)", but one-third of the data for this variable was missing (classified as Not Known). The service was unable to recall the exact circumstances under which this category was recorded at the time; this was partly due to the anonymized data and partly due to service restructuring that had taken place since this study was completed. Therefore, it remains unclear why these data were missing but, for whatever reason, it has undermined the reliability of the results.

The clinical outcomes were examined in two ways, taking into account both symptom scores and clinical status. The definition of recovery was based on previous evaluation studies (e.g. NEPHO, 2010), facilitating the comparison between the current results and the national data. However, the clinical status of recovery in this study was defined by cut-off points on questionnaires; future studies should include clinician diagnosed clinical status as an outcome measure. Furthermore, it should be noted that there are other ways to define recovery and clinical outcomes, such as quality of life and relapse rates (Kuyken et al., 2008). Future studies should consider broadening the range of outcome measures. It would also be helpful to compare the cost-effectiveness, which was not possible in this study due to the unavailability of the cost data.

Implications for service delivery

As mentioned in the Introduction, there has been a considerable variability in the clinical outcomes amongst the IAPT sites (NEPHO, 2010). This study revealed encouraging data with regards to the effectiveness of the treatments provided at the IAPT service in Suffolk. This service demonstrated the use of NICE recommended and evidence-based interventions (e.g. 86% of those in the high intensity group received CBT), supported by a considerable proportion of experienced therapists. These two factors have been previously associated with better clinical outcomes (Gyani et al., 2011). The fact that the two groups matched in their outcomes further suggested that low intensity interventions are equally effective, despite their being shorter in duration and delivered by less experienced therapists. This was in line with the assumption underlying the stepped care model (Bower and Gilbody, 2005). Ongoing data collection, as in this exercise, is helpful in monitoring progress, identifying the strengths and weaknesses of the services and unpacking the underlying factors.

Future research

Although the outcomes were comparable with the national average and established effectiveness of CBT, it appeared that our existing services could at best achieve a recovery rate of about 50%. Research is needed to further our understanding of why some people do not benefit from treatments as well as others. This is particularly important to inform future service delivery, given that IAPT has ongoing plans to expand the services to include children and young people.

Conclusion

IAPT is still in an early stage of development and yet it has demonstrated encouraging treatment effectiveness. This study highlighted the service use and outcomes in the Suffolk service, and in particular compared clients who received high and low intensity interventions. Despite various methodological limitations, results showed that the service as a whole achieved high recovery rates. Further research building upon the current study in unpacking the relative strengths and weaknesses for high and low intensity treatments would be beneficial for future service delivery.

Acknowledgements

The authors would like to thank the Suffolk IAPT Service, which has been renamed as the Suffolk Wellbeing Service since the completion of the study, and in particular Dr Tuija Juusti-Butler for providing helpful information.

References

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders DSM-IV (4th ed.)*. Washington, DC: Author.

Bower, P. and Gilbody, S. (2005). Stepped care in psychological therapies: access, effectiveness and efficiency. Narrative literature review. *British Journal of Psychiatry*, *186*, 11–17.

Clark, D. M. and Ehlers, A. (1993). An overview of the cognitive theory and treatment of panic disorder. *Applied and Preventative Psychology*, *2*, 131–139.

Clark, D. M., Layard, R., Smithies, R., Richards, D. A., Suckling, R. and Wright, B. (2009). Improving access to psychological therapy: initial evaluation of the two demonstration sites. *Behaviour Research and Therapy*, *47*, 910–920.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*. New York: Academic Press.

Gyani, A., Shafran, R., Layard, R. and Clark, D. M. (2011). Enhancing recovering rates in IAPT services: lessons from analysis of the year one data. Retrieved from http://www.iapt.nhs.uk.

Howell, D. C. (2002). *Statistical Methods for Psychology (5th ed.)*. Pacific Grove, CA: Duxbury.

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

Kuyken, W., Byford, S., Taylor, R. S., Watkins, E., Holden, E., White, K., et al. (2008). Mindfulness based cognitive therapy to prevent relapse in recurrent depression. *Journal of Consulting and Clinical Psychology*, *76*, 966–978.

Layard, R., Bell, S., Clark, D. M., Knapp, M., Meacher, M., Priebe, S., et al. (2006). *The Depression Report: a new deal for depression and anxiety disorders*. London: Centre for Economic Performance, London School of Economics.

National Institute for Health and Clinical Excellence (2004a). *Depression: management of depression in primary and secondary care (CG23)*. Retrieved from http://guidance.nice.org.uk/CG23.

National Institute for Health and Clinical Excellence (2004b). *Generalised Anxiety Disorder and Panic Disorder (with or without Agoraphobia) in Adults (CG113)*. Retrieved from http://guidance.nice.org.uk/CG113.

North East Public Health Observatory (NEPHO) (2010). *Improving Access to Psychological Therapies: a review of the progress by the sites in the first roll-out year*. Stockton on Tees: NEPHO.

Scogin, F., Hanson, A. and Welsh, D. (2003). Self-administered treatment in stepped-care models of depression treatment. *Journal of Clinical Psychology*, *59*, 341–349.

Seekles, W., van Straten, A., Beekman, A., van Marwijk, H. and Cuijpers, P. (2011). Stepped care treatment for depression and anxiety in primary care: a randomised controlled trial. *Trials*, *12*, 171.

Spitzer, R. L., Kroenke, K., Williams, J. B. W. and Lowe, B. (2006). A brief measure for assessing generalised anxiety disorder: the GAD-7. *Archives of Internal Medicine*, *166*, 1092–1097.

Shapiro, D. A., Barkham, M., Rees, A., Hardy, G. E., Reynolds, S. and Startup, M. (1994). Effects of treatment duration and severity of depression on the effectiveness of cognitive-behavioural and psychodynamic-interpersonal psychotherapies. *Journal of Consulting and Clinical Psychology*, *62*, 522–534.

van Straten, A., Tiemens, B., Hakkaart, L., Nolen, W. A. and Donker, M. C. H. (2006). Stepped care vs. matched care for mood and anxiety disorders: a randomized trial in routine practice. *Acta Psychiatrica Scandinavica*, *113*, 468–476.