Demographic and clinical predictors of response to internet-enabled cognitive–behavioural therapy for depression and anxiety

Ana Catarino, Sarah Bateup, Valentin Tablan, Katherine Innes, Stephen Freer, Andy Richards, Richard Stott, Steven D. Hollon, Samuel Robin Chamberlain, Ann Hayes and Andrew D. Blackwell

Background

Common mental health problems affect a quarter of the population. Online cognitive–behavioural therapy (CBT) is increasingly used, but the factors modulating response to this treatment modality remain unclear.

Aims

This study aims to explore the demographic and clinical predictors of response to one-to-one CBT delivered via the internet.

Method

Real-world clinical outcomes data were collected from 2211 NHS England patients completing a course of CBT delivered by a trained clinician via the internet. Logistic regression analyses were performed using patient and service variables to identify significant predictors of response to treatment.

Results

Multiple patient variables were significantly associated with positive response to treatment including older age, absence of long-term physical comorbidities and lower symptom severity at start of treatment. Service variables associated with positive response to treatment included shorter waiting times for initial assessment and longer treatment durations in terms of the number of sessions.

Conclusions

Knowledge of which patient and service variables are associated with good clinical outcomes can be used to develop personalised treatment programmes, as part of a quality improvement cycle aiming to drive up standards in mental healthcare. This study exemplifies translational research put into practice and deployed at scale in the National Health Service, demonstrating the value of technology-enabled treatment delivery not only in facilitating access to care, but in enabling accelerated data capture for clinical research purposes.

Common mental health disorders including depression and anxiety are characterised by intense emotional distress and have an impact on social and occupational functioning. One in four adults are estimated to have a mental health problem in any given year[1,2] but a significant proportion remain undiagnosed or untreated.[3] The Improving Access to Psychological Therapies (IAPT) programme is a large-scale initiative aimed at increasing access to accountable, evidence-based psychological therapy for common mental health disorders within the English National Health Service (NHS), while controlling costs.[4] In 2015/2016 one-third of patients referred to IAPT received cognitive–behavioural therapy (CBT).[5] With proven clinical effectiveness, structured CBT models are also amenable to delivery via 'self-help' programmes including online materials,[6,7] with various online CBT models shown as efficacious in controlled trial settings.[6–8] Online CBT offers a flexibility that is not possible under face-to-face programmes, allowing patients to undergo treatment outside of office hours, optimising convenience and avoiding perceived stigma associated with undertaking in-person therapy.[9] Other potential benefits include shorter waiting times, enhanced disclosure and improved access for patients who are reluctant to contact services given the nature of their condition (for example agoraphobia), or cannot travel because of disability or geographical location.[9,10] Despite multiple potential advantages, varying degrees of therapist intervention in online CBT can affect clinical outcomes and patient engagement in real-world settings,[5,8,11,12] with therapist-guided online CBT associated with better clinical outcomes and lower drop-out rates than self-guided programmes. Internet-enabled CBT (IECBT) is an effective type of online therapy,[13] successfully used within IAPT, whereby patients are offered weekly one-to-one sessions with an accredited CBT therapist. As a result of its one-to-one nature, IECBT is classed as a high-intensity therapy and can be used to treat more severe patients, relative to other self-guided and guided self-help online CBT modalities that are classed as low-intensity interventions and therefore only suitable for patients with milder presentations.[14] Previous research investigating predictors of clinical outcomes for low-intensity guided self-help interventions shows that higher levels of adherence to treatment and treatment credibility are associated with higher rates of improvement and lower post-treatment scores.[15,16] This highlights the importance of investigating predictors of clinical outcomes in response to high-intensity online interventions like IECBT, where the synchronous, yet anonymous, nature of the interaction between therapist and patient may promote treatment credibility and patient adherence to treatment protocol. The present study aims to explore patient and service variables as predictors of clinical outcomes, in patients receiving IECBT for the treatment of a common mental health disorder.

Method

Data were analysed from patients receiving IECBT for the treatment of a mental health disorder, between April 2015 and March 2016. IECBT was delivered using a commercial package, originally developed for and currently used in the English NHS, provided by Ieso Digital Health (http://uk.iesohealth.com). Patients self-referred or were referred by a primary healthcare worker directly to the service in the regions of Surrey, West Kent, Camden and East Riding of Yorkshire. Patients can register with the service using an online registration form or over the phone. Patients reporting suicidal intent, during registration or at any point during the episode of care, were appropriately advised online by their therapist or another member of the clinical team, and signposted to specialist services accordingly. In exceptional circumstances of immediate or serious risk patients were contacted over the phone by their therapist or a clinical supervisor.

After registration, patients were assigned to a qualified CBT therapist accredited by the British Association for Behavioural and Cognitive Psychotherapies. Initial assessments were carried out in an online therapy room via one-to-one real-time written conversation, after which the therapist assigned the patient a diagnosis and National Institute for Health and Care Excellence-approved disorder-specific CBT treatment protocols,[14] based on Roth & Pilling's CBT competences framework,[17] were delivered during weekly sessions. Treatment duration was determined by the therapist based on their clinical judgement, with typical treatment length between 6 and 8 sessions. Between treatment appointments, asynchronous messages and homework tasks were exchanged between therapist and patient, promoting engagement and adherence to evidence-based treatment models. All communication between therapist and patient was done exclusively online through Ieso's proprietary platform, following internationally recognised standards for information security (ISO 27001; https://www.iesohealth.com/en-gb/legal/iso-certificates).

Clinical outcomes in the IECBT group were referenced against reported outcomes for patients referred to IAPT services in the same time period and same regions where IECBT was offered. Patients in the reference group received care as usual, comprising high- and low-intensity treatments, face-to-face and online therapy services, including IECBT.[18]

The information captured through IAPT's minimum data-set, including IECBT, is intended to support monitoring of implementation and effectiveness of national policy/legislation, policy development, performance analysis and benchmarking, national analysis and statistics and national audit of IAPT services. At registration patients agree to the services' terms and conditions, including use of anonymised data for audit purposes and to support research, including academic publications or conference presentations.

Outcomes measures

Clinical outcomes were measured in terms of clinical recovery and improvement, defined following IAPT guidelines.[4,19] According to IAPT convention, these measures are defined for patients undergoing a minimum of two sessions of therapy. This is the minimum dose of therapy a patient must receive such that pre- and post-treatment scores are collected and clinical change can be estimated.[19] Clinical recovery and improvement are calculated based on two severity measures completed by the patient at initial assessment and for every therapy session (completion rate 95%): Patient Health Questionnaire (PHQ)-9[20] and General Anxiety Disorder (GAD)-7,[21] corresponding to depressive and anxiety symptoms, respectively.

The PHQ-9 is a nine-item measure designed to facilitate screening and severity assessment of depression, ranging from 0 to 27 and with a recommended cut-off of ten or more for distinguishing patients considered to have clinically significant depressive symptoms.[20] A reduction of six points or more on the PHQ-9 scale between two time points is indicative of statistically reliable improvement in symptom severity.[20]

The GAD-7 is a seven-item screening and severity measure for generalised anxiety disorder, ranging from 0 to 21 and with a recommended cut-off of eight or more for distinguishing patients considered to have clinically significant anxiety symptoms.[21,22] A reduction of four points or more on the GAD-7 scale between two time points is indicative of statistically reliable improvement in symptom severity.[21]

If a patient scores above the clinical threshold for one or both of these measures at initial assessment (i.e. ten or above for PHQ-9 and/or eight or above for GAD-7), they are classed as meeting 'caseness' at assessment. Other symptom severity measures, such as severity scores for subtypes of anxiety disorders, were not examined as only PHQ-9 and GAD-7 are mandatorily collected within the IAPT framework.

For patients undergoing two or more therapy sessions, the difference between scores at initial assessment and last treatment session for PHQ-9 and GAD-7 is used to determine patients' recovery status. Recovery is a binary measure. Under IAPT guidelines, patients with two or more therapy sessions who move from above caseness at assessment to below caseness on both the PHQ-9 and GAD-7 scales at the last treatment session are classed as recovered. The recovery rate for a group of patients is calculated as number of patients recovered, divided by number of patients at caseness at initial assessment.

Improvement is also a binary measure. Under IAPT guidance patients with two or more therapy sessions who show a significant reduction in at least one of the outcome measures from assessment to the last treatment session, while not showing a significant increase in the other outcome measure, were classed as improved (i.e. decrease of six points or more in the PHQ-9 and/or four points or more in the GAD-7, while not simultaneously showing an increase of six points or more in the PHQ-9 or four points or more in the GAD-7). The improvement rate for a group of patients is calculated as number of patients improved, divided by number of patients with two or more therapy sessions. Patients who simultaneously improve and recover are classed as reliably recovered.

Sample size

A total of 4468 patients registered with the IECBT service between April 2015 and March 2016. Of these, 487 patients (11%) did not meet the eligibility criteria (over 18 years old, registered with a general practitioner in the geographical region where the service is commissioned) and were signposted to other mental health services as appropriate. From the remaining 3981 eligible patients, 95 (2%) were deemed not suitable for the service for clinical reasons (for example risk, Axis II disorder) and were signposted to other mental health services as appropriate. A total of 3886 patients were offered treatment, of which 2211 (57%) had two or more treatment sessions. Out of these 2211 patients, 1818 (82%) were at caseness at assessment (170 at caseness according to PHQ-9 only, 375 at caseness for GAD-7 only and 1273 at caseness for both) (Fig. 1). A comparison of demographics between patients undergoing two or more therapy sessions and patients who dropout before this point can be found in the supplementary materials (Supplementary Table 1, available at https://doi.org/10.1192/bjo.2018.57).

Between April 2015 and March 2016 a total of 45 560 referrals were received by IAPT services in the same regions where IECBT was offered. In total, 19 325 patients were discharged in this time
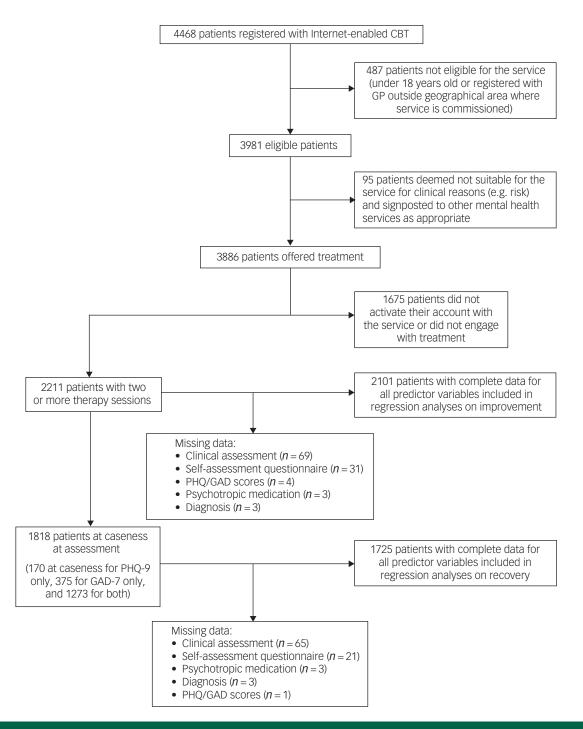
Fig. 1 Study profile and patient flow chart.

GP, general practitioner; PHQ, Patient Health Questionnaire; GAD General Anxiety Disorder.

period having had two therapy sessions or more, of which 17 470 (90%) were at caseness at assessment. Data analyses focusing on the improvement metric were conducted on data from patients with two or more therapy sessions, while analyses focusing on the recovery metric were conducted on data from patients at caseness at assessment, who also had two or more therapy sessions.

Regression analyses – predictors of clinical response in IECBT

Logistic regression analyses were performed in R to identify significant predictors of recovery and improvement in patients receiving IECBT, based on patient demographics and service variables. Demographic variables included patient gender, age, severity, diagnosis, whether or not the patient had a long-term physical condition, and whether or not the patient was taking psychotropic medication (for example antidepressants or anxiolytics) at the start of treatment. Service variables comprised data pertaining to a patient's course of treatment, including waiting times between various stages in the patient journey, treatment duration and number of scheduled appointments the patient failed to attend.

Given the nature of the statistical models employed, record sets were included only for cases with complete data for all predictor variables. Of the initial sample of 2211 patients with two therapy sessions and 1818 patients at caseness at assessment, 95% had

complete data for all predictor variables and were included in the analyses ($n = 2101$ for improvement analysis, $n = 1725$ for recovery analysis; Fig. 1).

Continuous predictor variables were scaled and centred to the mean. Multicollinearity analyses were performed to investigate potential correlations between predictor variables. Statistical significance was defined as $P<0.05$ two-tailed, uncorrected.

Comparative analysis of clinical outcomes

Although inferential analysis of comparative clinical effectiveness is not possible in the present study because of the lack of a face-to-face control group, publicly available IAPT data makes it possible to reference IECBT clinical outcomes against averages for the same time period and geographical regions. Patients with two or more therapy sessions in the IECBT group ($n = 2211$) were matched to the IAPT reference group ($n = 19 325$) for severity (PHQ-9 and GAD-7 scores at assessment), using a multivariate rejection sampling algorithm implemented in R.[23,24] Lack of publicly available distribution data for other variables means it was not possible to match the two groups for other potentially relevant variables such as age, diagnosis and presence of long-term physical comorbidities. Clinical outcomes of the severity-matched IECBT group relative to IAPT are reported.

Results

Regression analyses – predictors of clinical response in IECBT

Logistic regression analyses identified presence of long-term physical conditions, initial GAD-7 scores, waiting time for assessment,

total number of treatment sessions and patient age as significant predictors of improvement (Table 1). Apart from waiting time for assessment, these variables were also identified as significant predictors of recovery, in addition to initial PHQ-9 scores (Table 2).

Results show that patients with long-term physical conditions are less likely to show good clinical outcomes compared with patients without long-term conditions (Tables 1 and 2). Patients with higher severity scores at assessment are also less likely to show clinical recovery (Table 2). However, in line with what can be observed from Table 1, results suggest that patients with higher GAD-7 scores at assessment have a higher likelihood of showing clinical improvement.

A significant positive association between patient age and likelihood of good clinical outcomes was also observed. This association was explored further in a *post hoc* analysis that revealed a significant negative correlation between patient age and severity (age and PHQ-9: $r = -0.09$, $t = -4.02$, d.f. = 2102, $P<0.001$; age and GAD-7: $r = -0.13$, $t = -5.98$, d.f. = 2102, $P<0.001$), as well as a weak but significant positive correlation between patient age and number of treatment sessions ($r = 0.05$, $t = 2.10$, d.f. = 2102, $P=0.036$).

Finally, results show that patients who have undergone a larger number of therapy sessions are more likely to show good clinical outcomes (Tables 1 and 2). However, *post hoc* analyses showed no significant association between treatment duration and clinical outcomes in patients with five or more sessions. Clinical outcome rates were optimal and less variable for treatment durations of five to nine sessions (51% of patients with more than two sessions, recovery rate: 57–60%; improvement rate: 67–72%). Clinical outcomes for patients with more than two but fewer than five treatment sessions are significantly lower and more variable (14% of all patients with

Table 1 Results of logistic regression analysis investigating predictors of improvement in the internet-enabled cognitive–behavioural therapy cohort ($n = 2101$)[a]

Predictor variable	Mean/prevalence	b	s.e.	Wald's statistic, z^2	P	Subgroup n	Improvement rate, %
Gender, %							
Men	27.3	–	–	–	–	574	60.1
Women	72.1	0.04	0.11	0.15	0.698	1514	62.2
Unknown/not stated	0.6	−0.24	0.59	0.17	0.682	13	53.8
Age, years: mean	36.8	0.12	0.05	5.87	0.015*	–	–
Long-term physical conditions, %							
No	34.3	–	–	–	–	721	60.9
Yes	20.8	−0.32	0.14	5.29	0.021*	436	59.9
Unknown/not stated	44.9	0.02	0.11	0.05	0.819	944	62.9
Diagnosis,[b] %							
Anxiety	42.1	–	–	–	–	885	61.4
Depression	22.6	0.03	0.13	0.04	0.851	474	60.1
Other	35.3	0.13	0.11	1.30	0.254	742	62.8
Psychotropic medication, %							
Prescribed not taking	8.1	–	–	–	–	171	59.6
Prescribed taking	38.3	0.24	0.18	1.78	0.182	805	63.6
Not prescribed	52.5	0.26	0.18	2.05	0.152	1103	60.4
Unknown/not stated	1.0	0.23	0.49	0.21	0.649	22	63.6
StartPhq9, mean	12.7	0.05	0.07	0.50	0.479	–	–
StartGad7, mean	11.9	0.53	0.06	70.91	<0.001***	–	–
WaitingSAQ, days: mean	3.0	0.01	0.06	0.02	0.894	–	–
WaitingAssignment, days: mean	0.8	−0.04	0.06	0.44	0.508	–	–
WaitingContact, days: mean	1.1	−0.07	0.08	0.93	0.336	–	–
WaitingAssessment, days: mean	7.6	−0.14	0.05	7.54	0.006**	–	–
WaitingTreatment, days: mean	8.9	−0.02	0.05	0.18	0.671	–	–
NumberSessions, sessions: mean	5.5	0.29	0.05	32.09	<0.001***	–	–
NumberDNA, sessions: mean	0.5	−0.01	0.05	0.06	0.813	–	–

StartPhq9, Patient Health Questionnaire-9 score at assessment; StartGad7, General Anxiety Disorder-7 score at assessment; Waiting SAQ, time between referral and patient completing self-assessment questionnaire; WaitingAssignment, waiting time from patient completing the self-assessment questionnaire and therapist assignment; WaitingContact, waiting time between therapist assignment and first contact by the therapist; WaitingAssessment, waiting time between first contact from the therapist and clinical assessment appointment; WaitingTreatment, waiting time between clinical assessment and first therapy session; NumberSessions, total number of therapy sessions attended by the patient; NumberDNA, number of scheduled appointments the patient failed to attend.
a. Men, long-term physical conditions 'no', diagnosis 'anxiety' and psychotropic medication 'prescribed not taking' were reference classes for the categorical variables.
b. The diagnosis anxiety encompasses patients diagnosed with agoraphobia, generalised anxiety disorder, hypochondriacal disorder, obsessive–compulsive disorder, panic disorder, post-traumatic stress disorder, social phobia, specific phobia or anxiety disorder unspecified. The diagnosis depression encompasses patients diagnosed with depressive episode, dysthymia or recurrent depressive disorder. The diagnosis 'other' encompasses all diagnoses not otherwise listed (for example chronic pain, bereavement, eating disorders).
*$P<0.05$, **$P<0.01$, ***$P<0.001$.

Table 2 Results of logistic regression analysis investigating predictors of recovery in the internet-enabled cognitive–behavioural therapy cohort ($n = 1725$)[a]

Predictor variable	Mean/prevalence	b	s.e.	Wald's statistic, z^2	P	Subgroup, n	Recovery rate, %
Gender, %							
Men	26.5	–	–	–	–	457	51.0
Women	73.0	0.15	0.12	1.62	0.202	1259	50.8
Unknown/not stated	0.5	–0.14	0.70	0.04	0.846	9	44.4
Age, years: mean	36.3	0.23	0.06	18.05	<0.001***	–	–
Long-term physical conditions, %							
No	33.2	–	–	–	–	573	55.8
Yes	22.0	–0.37	0.15	5.88	0.015*	379	43.5
Unknown/not stated	44.8	–0.14	0.12	1.32	0.251	773	50.6
Diagnosis, %[b]							
Anxiety	40.8	–	–	–	–	703	53.3
Depression	23.6	–0.04	0.15	0.07	0.796	407	45.9
Other	35.7	0.02	0.13	0.02	0.875	615	51.1
Psychotropic medication, %							
Prescribed not taking	8.8	–	–	–	–	151	47.0
Prescribed taking	41.7	0.01	0.20	0.01	0.942	719	46.3
Not prescribed	48.3	0.06	0.20	0.09	0.769	834	55.6
Unknown/not stated	1.2	–0.49	0.53	0.84	0.360	21	38.1
StartPhq9, mean	14.3	–0.56	0.07	70.32	<0.001***	–	–
StartGad7, mean	13.4	–0.30	0.06	24.76	<0.001***	–	–
WaitingSAQ, days: mean	3.0	0.01	0.06	0.02	0.900	–	–
WaitingAssignment, days: mean	0.8	0.04	0.06	0.36	0.550	–	–
WaitingContact, days: mean	1.2	–0.11	0.07	2.16	0.142	–	–
WaitingAssessment, days: mean	7.4	–0.08	0.06	2.16	0.141	–	–
WaitingTreatment, days: mean	8.8	–0.02	0.05	0.18	0.674	–	–
NumberSessions, sessions: mean	5.6	0.32	0.05	33.83	<0.001***	–	–
NumberDNA, sessions: mean	0.5	–0.11	0.05	3.81	0.051	–	–

StartPhq9, Patient Health Questionnaire-9 score at assessment; StartGad7, General Anxiety Disorder-7 score at assessment; Waiting SAQ, time between referral and patient completing self-assessment questionnaire; WaitingAssignment, waiting time from patient completing the self-assessment questionnaire and therapist assignment; WaitingContact, waiting time between therapist assignment and first contact by the therapist; WaitingAssessment, waiting time between first contact from the therapist and clinical assessment appointment; WaitingTreatment, waiting time between clinical assessment and first therapy session; NumberSessions, total number of therapy sessions attended by the patient; NumberDNA, number of scheduled appointments the patient failed to attend.
a. Men, long-term physical conditions 'no', diagnosis 'anxiety' and psychotropic medication 'prescribed not taking' were reference classes for the categorical variables.
b. The diagnosis anxiety encompasses patients diagnosed with agoraphobia, generalised anxiety disorder, hypochondriacal disorder, obsessive–compulsive disorder, panic disorder, post-traumatic stress disorder, social phobia, specific phobia or anxiety disorder unspecified. The diagnosis depression encompasses patients diagnosed with depressive episode, dysthymia or recurrent depressive disorder. The diagnosis 'other' encompasses all diagnoses not otherwise listed (for example chronic pain, bereavement, eating disorders).
*P<0.05, ***P<0.001.

more than two sessions, recovery rate: 27–53%; improvement rate: 42–61%).

Tests of the full models against constant-only models were significant for both regression analyses (improvement regression model: $\chi^2(19) = 195.95$, $P<0.001$; recovery regression model: $\chi^2(19) = 278.36$, $P<0.001$). Additionally, multicollinearity analyses revealed variance inflation factors smaller than two for all predictor variables. This is the standard threshold value for indicating the presence of multicollinearity in this type of analysis, thus confirming that regression models were not affected by the presence of multicollinearity.[25]

IAPT's improvement and recovery metrics are, by definition, biased by initial symptom severity, which confounds interpretation of the results. In a *post hoc* regression analysis investigating predictors of clinical outcomes while controlling for artefactual relations with initial severity, we defined per cent improvement as a 25% decrease in scores for both the PHQ-9 and the GAD-7.[26] Similar to IAPT convention, we classed a patient as achieving per cent improvement if they showed a 25% decrease in scores in one or both scales, without showing deterioration in either scale. Results of this analysis show broadly similar results to the analysis on predictors of improvement as defined according to IAPT convention, but they no longer show the significant association between initial severity scores and per cent improvement (Supplementary Table 2).

Following IAPT convention, improvement, recovery and per cent improvement metrics were defined combining the PHQ-9 and GAD-7 scales. Although this allows for a more comprehensive characterisation of the patients, who often present with a combination of depressive and anxiety features, these are two separate scales, measuring different constructs. It can be hypothesised that patient and service variables may have an impact on the likelihood of good clinical outcomes for each scale differently. *Post hoc* regression analyses investigating predictors of per cent improvement for each scale separately are presented in Supplementary Tables 3 and 4.

Clinical outcomes adjusted for symptom severity and benchmarked against national audit comparator data

Variations were observed in the likelihood of improvement and recovery with PHQ-9 and GAD-7 scores at assessment, for patients treated with IECBT. Regression analyses results show that patients with higher PHQ-9 and GAD-7 scores at assessment, have lower likelihoods of recovery but equivalent or higher likelihood for improvement. Accordingly, IECBT clinical improvement and recovery data were benchmarked against severity-matched cohorts. Severity-matched patients treated with IECBT showed similar improvement and recovery rates relative to IAPT patients, as well as similar magnitude of symptom reduction, pre- and post-treatment (Table 3). Although classical significance testing was avoided because of bias in favour of rejecting the null hypothesis for large sample sizes,[27] effect sizes and 95% confidence intervals are presented (Table 3). Despite some isolated differences in disorder distribution across the two cohorts, the observed odds ratios and effect sizes are generally small for most variables.[28] Together with differences in clinical outcomes of less than 1% and differences in magnitude of symptom reduction of 0.6 points or less between the two groups, these results suggest that differences between the

Table 3 Demographic details and clinical outcomes for patients finishing a course of treatment between April 2015 and March 2016[a]

	IECBT (severity matched) (n = 2211)		IAPT (areas where IECBT was offered) (n = 19 325)		Group difference	
	Mean (s.d.) 95% CI	% (95% CI)	Mean (s.d.) 95% CI	%, 95% CI	Cohen's d (95% CI)	OR (95% CI)
Gender, women	–	73 (71–75)	–	67 (67–68)	–	1.30 (1.18 to 1.44)
Long-term physical conditions	–	21 (20–23)	–	23 (22–23)	–	0.92 (0.83 to 1.03)
Diagnosis[b]						
Depression	–	27 (25–29)	–	36 (35–36)	–	0.68 (0.62 to 0.75)
Anxiety	–	41 (39–43)	–	38 (37–39)	–	1.13 (1.03 to 1.24)
Other	–	32 (30–34)	–	27 (26–27)	–	1.30 (1.18 to 1.43)
PHQ-9 at assessment	13.8 (4.5) 13.6 to 14.0	–	14.3 (5.2) 14.2 to 14.4	–	−0.10 (−0.14 to −0.05)	–
GAD-7 at assessment	13.0 (3.8) 12.8 to 13.1	–	13.2 (4.2) 13.1 to 13.3	–	−0.05 (−0.09 to 0.004)	–
PHQ-9 post-treatment	8.6 (6.1) 8.3 to 8.9	–	8.5 (6.2) 8.4 to 8.6	–	0.02 (−0.03 to 0.06)	–
GAD-7 post-treatment	8.0 (5.4) 7.8 to 8.2	–	7.8 (5.4) 7.7 to 7.9	–	0.04 (−0.01 to 0.08)	–
Treatment length, sessions	5.5 (2.8) 5.4 to 5.6	–	6.4[c]	–	N/A	–
Improvement rate	–	65.8 (63.8 to 67.8)	–	66.2 (65.5 to 66.8)	–	0.98 (0.90 to 1.08)
Recovery rate	–	48.8 (46.7 to 51.0)	–	49.5 (48.8 to 50.3)	–	0.97 (0.89 to 1.06)

CBT, cognitive–behavioural therapy; IECBT, internet-enabled CBT; IAPT, Improving Access to Psychological Therapies; PHQ, Patient Health Questionnaire; GAD, General Anxiety Disorder; NA, not available.
a. The IECBT cohort was severity matched to the IAPT cohort using a multivariate rejection sampling.
b. The diagnosis anxiety encompasses patients diagnosed with agoraphobia, generalised anxiety disorder, obsessive–compulsive disorder, hypochondriacal disorder, panic disorder, post-traumatic stress disorder, social phobia, specific phobia or anxiety disorder unspecified. The diagnosis depression encompasses patients diagnosed with depressive episode, dysthymia or recurrent depressive disorder. The diagnosis 'other' encompasses all diagnoses not otherwise listed (for example chronic pain, bereavement, eating disorders).
c. Average treatment duration for all conditions nationwide, regional data not available.

two groups in clinical outcomes and score reduction are unlikely to be meaningful.[29]

Discussion

This paper reports the first real-world (non-randomised controlled trial) clinical outcomes data for patients with depression and anxiety treated using internet-enabled CBT. The first evidence of clinical efficacy of IECBT in depression was published in the Lancet in 2009.[13] These extended data, including both depression and anxiety disorders, offer an example of translational research put into practice and successfully deployed at scale. The application of the resultant data-set in advancing our understanding of clinical and demographic variables associated with response to treatment suggests that there might be value in data-enabled mental health services as platforms for clinical research. Knowledge acquired with these tools can be used to refine service specifications and develop personalised treatment programmes, as part of a quality improvement cycle aiming to drive up standards in mental healthcare.

Main findings

Regression analyses revealed a significant association between initial psychometric scores and likelihood of recovery, with higher scores associated with lower recovery rates. By definition, patients recover by going below the caseness threshold for both PHQ-9 and GAD-7. Therefore, it is not unexpected that patients whose initial scores are closer to that threshold have higher chances of recovery. This does, however, raise the question of whether recovery alone is a suitable index to measure clinical outcome.

As indicated by Gyani and colleagues[19] the recovery metric does not take into account whether the observed reduction in severity is greater than the measurement error of the scales. Conversely the improvement index is a measure of whether or not a reduction in severity is statistically reliable, regardless of caseness and may therefore be a better metric for widespread use. In the present study, patients with higher initial scores are more likely to show clinical improvement, as validated by the results of the regression analysis, where patients with higher initial GAD-7 scores show higher likelihoods of improvement (Table 1). IAPT's reliable recovery index is a composite metric measuring whether a patient recovered while simultaneously showing a statistically reliable reduction in severity. Although by definition this metric may be less susceptible to bias in favour of patients who are near the recovery threshold, it will still be biased against patients with higher severity scores at assessment, who will be less likely to cross the recovery threshold.

To investigate predictors of clinical outcomes while controlling for artefactual relations with initial severity, we conducted a *post hoc* analysis investigating predictors of per cent improvement.[26] Results show broadly similar results to the regression analysis on predictors of improvement as defined according to IAPT convention, but the significant association with initial severity scores is no longer present with the per cent improvement measure (Supplementary Table 2). Although differences in recovery rate with severity may be expected in this context, they may also indicate the presence of non-specific treatment effects. Future strategies to improve treatment effectiveness should therefore be aimed at boosting recovery of more severe patients, including increased session frequency at the start of treatment, or the use of specific CBT protocols for severe depression.

Regression analyses on improvement and recovery also revealed significant associations between clinical outcomes and age, presence of a long-term physical condition and number of therapy sessions. Results show that greater age is associated with better clinical

outcomes, in contrast with previous research showing lower effectiveness of CBT in older adults.[30] However, it is important to note that in the present study the mean age of the patient cohort was 36 years, whereas previous research on the effects of CBT on older adults focused on adults over the age of 55.[30] Older adults are more likely to be affected by age-related cognitive decline and physical comorbidities that may directly influence CBT outcomes but are not prevalent factors in the current cohort.

Post hoc analyses on predictors of per cent improvement reveal age to be a positive predictor of likelihood of per cent improvement, similar to what was observed for the analysis on improvement defined under IAPT's convention. This suggests that despite a significant negative correlation between patient age and severity, the association between age and clinical outcomes is not in this case driven by differences in severity across the age range in this particular cohort. A weak but significant positive correlation between patient age and number of treatment sessions, as well as higher mean age of patients with two or more therapy sessions (Supplementary Table 1) suggests that in this particular cohort, older patients may be less likely to drop-out at earliest stages of treatment, therefore benefitting from a larger therapeutic dose and consequently be more likely to achieve positive clinical outcomes.

In this cohort it was also observed that patients with long-term physical conditions were less likely to show good clinical outcomes. This finding is unsurprising given that long-term physical conditions are often associated with comorbid mental health problems and complex psychological issues, which may themselves be chronic in nature and often treatment resistant.[31,32] Lower probability of response to treatment may signal the need for tailored, condition-specific CBT models, so patients can be helped to deal with mental and physical symptoms in an integrated fashion. However, it could also be argued that PHQ-9 and GAD-7, used to calculate clinical outcomes, lack sensitivity to detect clinically significant improvements in patients with long-term conditions. Disease-specific measures that may better reflect the complexities of the physical disease and provide better indicators of clinical improvement in these patients, should be included in future studies.

Service variables shown to be associated with the likelihood of good clinical outcomes included higher number of therapy sessions and reduced waiting time for assessment. Although these findings are supported by similar reports in the literature,[33] care should be taken when drawing causal inferences. At first glance these results convey the impression that longer courses of treatment are associated with better clinical outcomes. However, an alternative explanation is that patients who do not adhere to their treatment plan and drop-out at an earlier stage during the course of treatment, therefore receiving a subtherapeutic dose, are less likely to achieve good clinical outcomes. This hypothesis is supported by results of *post hoc* analyses showing that clinical outcome rates were optimal and less variable for treatment durations of five to nine sessions, whereas patients with more than two but fewer than five treatment sessions observed significantly lower and more variable clinical outcomes. Difficulties with engagement leading to poor clinical outcomes may be particularly relevant in patients with more severe depressive symptoms, who by the nature of their condition may lack motivation to attend treatment sessions and generally adhere to their treatment plan.

Limitations

A numerical comparison of IECBT clinical outcomes against IAPT's averages in the present study suggests that IECBT is as effective as standard care. The comparison between these two groups is presented to demonstrate general equivalence of IECBT and IAPT services, building on previous results from a clinical trial of IECBT[13] and supporting the effectiveness of this therapy modality in a real-world clinical setting. However, it is important to note there are several limitations for this analysis and caution should be taken not to overinterpret these findings. First, as this was an audit study and not a randomised controlled trial, group comparisons between patients receiving IECBT and IAPT patients are open to the effects of selection bias.

Second, although the IECBT group was matched to the reference group for severity, the aggregated nature of the data published in IAPT's annual reports means that it was not possible to use propensity analyses or selection algorithms to better match the patients who got IECBT to that subset of the patients in IAPT who were most similar to them.

Third, while all patients in the IECBT group received CBT, patients in the IAPT reference group received a range of different therapy types, including IECBT. IECBT is not suitable for all patients, including those at risk and those who are not literate, not fluent English speakers or who do not have access to an internet-connected device. It can be hypothesised that differences in therapy type, together with potential cohort differences in other uncontrolled variables such as presence of secondary comorbid mental health conditions, IQ and socioeconomic status, may also account for variance in clinical outcomes. Furthermore, while there is a large evidence base supporting the efficacy of CBT in general, the issue of side-effects and potential increase in suicidality as a result of therapy remains largely unexplored, for both face-to-face and online CBT modalities. The data collected in the present study and published in IAPT's annual reports did not allow us to explore possible side-effects of IECBT and other types of therapy offered under the IAPT programme.

Regarding predictive analyses in the present study, although these comprehensively included patient and service variables, it should be noted that therapy variables were not included. One advantage of IECBT's unique method, is that it provides full therapy session transcripts. Future work could make use of these data, together with advanced analysis techniques, such as natural language processing and machine learning, to identify therapeutic interventions that lead to the best proximal and distal clinical outcomes in patients.

Implications and future research

A positive aspect of the in-service data collection method used here and in other IAPT services is that replication of these findings is possible in a way that is often cost-prohibitive in clinical trials. Analysis of subsequent cohorts can inform whether these observations generalise to other cohorts and also add to the scientific knowledge of effective CBT change mechanisms. Understanding predictors of good clinical outcomes may facilitate development of improved, patient-focused, stratified/stepped-care allocation models and also enable the development of enhanced therapeutic protocols. Data derived from an outcomes-measurement framework is also of potential value to providers, who can adapt their services to better meet the needs of their patients and consistently monitor service quality and encourage accountability.

In England, continuous monitoring of clinical effectiveness using an outcomes-measurement framework has enabled systematic improvements in the quality and consistency of care delivery and enabled a transition from fee-for-service to fee-for-value payment models.[34] In the USA, as the Centres for Medicare and Medicaid Services begin to implement value-based payment models the importance of understanding what and why treatments work, and what their clinical and economic impact is, becomes evident. Translating the IAPT model, including digital approaches,

not only into the USA but worldwide could have a dual advantage, improving the quality and accountability of mental healthcare while reducing cost by enabling a shift towards capitated and fee-for-value payment models.

Ana Catarino, PhD, Senior Scientist, Clinical Science Laboratory at Ieso, Ieso Digital Health, UK; **Sarah Bateup**, BSc MEd, Chief Clinical Officer, Clinical Science Laboratory at Ieso, Ieso Digital Health, UK; **Valentin Tablan**, PhD, Senior Vice President for Artifical intelligence, Clinical Science Laboratory at Ieso, Ieso Digital Health, UK; **Katherine Innes**, BSc, Clinical Informatics & Visualisation Specialist, Clinical Science Laboratory at Ieso, Ieso Digital Health, UK; **Stephen Freer**, BA BSc PGDip, UK Clinical Lead, Clinical Science Laboratory at Ieso, Ieso Digital Health, UK; **Andy Richards**, PhD, Director, Ieso Digital Health, UK; **Richard Stott**, DClinPsy, Visiting Associate, Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK; **Steven D. Hollon**, PhD, Professor of Psychology, Department of Psychology, Vanderbilt University, USA; **Samuel Robin Chamberlain**, PhD, MRCPsych, Wellcome Trust Clinical Fellow and Honorary Consultant, Department of Psychiatry, University of Cambridge and Cambridge and Peterborough NHS Foundation Trust, UK; **Ann Hayes**, PhD, Independent Pharmaceutical and Digital Health Consultant, Ieso Digital Health, UK; **Andrew D. Blackwell**, PhD, Chief Scientific Officer, Clinical Science Laboratory at Ieso, Ieso Digital Health, UK

Correspondence: Ana Catarino, Clinical Science Laboratory at Ieso, Ieso Digital Health, The Jeffreys Building, Cowley Road, Cambridge CB4 0DS, UK. Email: a.catarino@iesohealth.com

First received 17 Apr 2018, final revision 17 Aug 2018, accepted 5 Sep 2018

Supplementary material

Supplementary material is available online at https://doi.org/10.1192/bjo.2018.57

Funding

References

1 Ferrari AJ, Charlson FJ, Norman RE, Patten SB, Freedman G, Murray CJL, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. *PLoS Med* 2013; **10**: e1001547.

2 McManus S, Bebbington P, Jenkins R, Brugha T. *Mental Health and Wellbeing in England*. Health and Social Care Information Centre, 2016.

3 McCrone P, Dhanasiri S, Patel A, Knapp M, Lawton-Smith S. Paying the price: the cost of mental health care in England to 2026. *Br J Psychiatry* 2008; **184**: 386–92.

4 Clark DM. Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: the IAPT experience. *Int Rev Psychiatry* 2011; **23**: 318–27.

5 Health and Social Care Information Centre. *Psychological Therapies: Annual Report on the Use of IAPT Services*. Health and Social Care Information Centre, 2016.

6 Olthuis JV, Watt MC, Bailey K, Hayden JA. Therapist-supported Internet cognitive behavioural therapy for anxiety disorders in adults. *Cochrane Database Syst Rev* 2015; **3**: CD011565.

7 Saddichha S, Al-Desouki M, Lamia A, Linden IA, Krausz M. Online interventions for depression and anxiety–a systematic review. *Heal Psychol Behav Med* 2014; **2**: 841–81.

8 Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. *Psychol Med* 2007; **37**: 319–28.

9 Gega L, Marks I, Mataix-Cols D. Computer-aided CBT self-help for anxiety and depressive disorders: experience of a London clinic and future directions. *J Clin Psychol* 2004; **60**: 147–57.

10 Beattie A, Shaw A, Kaur S, Kessler D. Primary-care patients' expectations and experiences of online cognitive behavioural therapy for depression: a qualitative study. *Health Expect* 2009; **12**: 45–59.

11 Andersson G, Cuijpers P. Pros and cons of online cognitive-behavioural therapy. *Br J Psychiatry* 2008; **193**: 270–1.

12 Dryman MT, McTeague LM, Olino TM, Heimberg RG. Evaluation of an open-access CBT-based Internet program for social anxiety: patterns of use, retention, and outcomes. *J Consult Clin Psychol* 2017; **85**: 988–99.

13 Kessler D, Lewis G, Kaur S, Wiles N, King M, Weich S, et al. Therapist-delivered internet psychotherapy for depression in primary care: a randomised controlled trial. *Lancet* 2009; **374**: 628–34.

14 National Institute for Health and Clinical Excellence. *Common Mental Health Disorders: the NICE Guideline on Identification and Pathways to Care*. The British Psychological Society and The Royal College of Psychiatrists, 2011.

15 El AS, Ljótsson B, Hedman E, Kaldo V, Andersson E, Rück C, et al. Predictors of symptomatic change and adherence in internet-based cognitive behaviour therapy for social anxiety disorder in routine psychiatric care. *PLoS One* 2015; **10**: e0124258.

16 Alaoui S E, Ljotsson B, Hedman E, Svanborg C, Kaldo V, Lindefors N. Predicting outcome in internet-based cognitive behaviour therapy for major depression: a large cohort study of adult patients in routine psychiatric care. *PLoS One* 2016; **11**: e0161191.

17 Roth AD, Pilling S. Accelerated publication using an evidence-based methodology to identify the competences required to deliver effective cognitive and behavioural therapy for depression and anxiety disorders. *Behav Cogn Psychother* 2008; **36**: 129–47.

18 Health and Social Care Information Centre. *IAPT Annual Report for 2015/2016*. NHS Digital, 2017 (http://content.digital.nhs.uk/iaptreports).

19 Gyani A, Shafran R, Layard R, Clark DM. Enhancing recovery rates: lessons from year one of IAPT. *Behav Res Ther* 2013; **51**: 597–606.

20 Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001; **16**: 606–13.

21 Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006; **166**: 1092–7.

22 Kroenke K, Spitzer RL, Williams JBW, Monahan PO, Lowe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Ann Intern Med* 2007; **146**: 317–25.

23 R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation, 2016 (http://www.r-project.org/).

24 Martino L, Míguez J. Generalized rejection sampling schemes and applications in signal processing. *Signal Process* 2010; **90**: 2981–95.

25 Myers R. *Classical and Modern Regression with Application (2nd edn)*. Duxbury/Thomson Learning, 1990.

26 Hiller W, Schindler AC, Lambert MJ. Defining response and remission in psychotherapy research: a comparison of the RCI and the method of percent improvement. *Psychother Res* 2012; **22**: 1–11.

27 Lin M, Lucas HC. Too big to fail: large samples and the p -value problem. *Inf Syst Res* 2013; **24**: 938–9.

28 Chen H, Cohen P, Chen S. How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Commun Stat Simul Comput* 2010; **39**: 860–4.

29 Tryon WW, Lewis C. An inferential confidence interval method of establishing statistical equivalence that corrects Tryon's (2001) reduction factor. *Psychol Methods* 2008; **13**: 272–7.

30 Gould RL, Coulson MC, Howard RJ. Efficacy of cognitive behavioral therapy for anxiety disorders in older people: a meta-analysis and meta-regression of randomized controlled trials. *J Am Geriatr Soc* 2012; **60**: 218–29.

31 Thorpe K, Jain S, Joski P. Prevalence and spending associated with patients who have a behavioral health disorder and other conditions. *Health Aff* 2017; **36**: 124–32.

32 Fenton WS, Stover ES. Mood disorders: Cardiovascular and diabetes comorbidity. *Curr Opin Psychiatry* 2006; **19**: 421–7.

33 Clark DM, Canvin L, Green J, Layard R, Pilling S, Janecka M. Transparency about the outcomes of mental health services (IAPT approach): an analysis of public data. *Lancet* 2018; **391**: 679–86.

34 NHS England and NHS Improvement. *Developing an Outcomes-Based Payment Approach for IAPT Services*. NHS Improvement, 2017 (https://improvement.nhs.uk/documents/661/IAPT_Payment_Guidance.pdf).