BMJ Open

A retrospective observational analysis to identify patient and treatment-related predictors of outcomes in a community mental health programme

Stuart A Green,[1] Emmi Honeybourne,[2] Sylvia R Chalkley,[1] Alan J Poots,[1] Thomas Woodcock,[1] Geraint Price,[2] Derek Bell,[1] John Green[2,3]

To cite: Green SA, Honeybourne E, Chalkley SR, *et al*. A retrospective observational analysis to identify patient and treatment-related predictors of outcomes in a community mental health programme. *BMJ Open* 2015;**5**:e006103. doi:10.1136/bmjopen-2014-006103

▸ Prepublication history for this paper is available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2014-006103).

Received 14 July 2014
Revised 31 March 2015
Accepted 9 April 2015

[1]NIHR CLAHRC Northwest London, Imperial College London, London, UK
[2]Central and North West London NHS Foundation Trust, London, UK
[3]Department of Clinical Health Psychology, St. Marys Hospital, London, UK

Correspondence to
Stuart A Green;
stuart.green@imperial.ac.uk

ABSTRACT

Objectives: This study aims to identify patient and treatment factors that affect clinical outcomes of community psychological therapy through the development of a predictive model using historic data from 2 services in London. In addition, the study aims to assess the completeness of data collection, explore how treatment outcomes are discriminated using current criteria for classifying recovery, and assess the feasibility and need for undertaking a future larger population analysis.

Design: Observational, retrospective discriminant analysis.

Setting: 2 London community mental health services that provide psychological therapies for common mental disorders including anxiety and depression.

Participants: A total of 7388 patients attended the services between February 2009 and May 2012, of which 4393 (59%) completed therapy, or there was an agreement to end therapy, and were included in the study.

Primary and secondary outcome measures: Different combinations of the clinical outcome scores for anxiety Generalised Anxiety Disorder-7 and depression Patient Health Questionnaire-9 were used to construct different treatment outcomes.

Results: The predictive models were able to assign a positive or negative clinical outcome to each patient based on 5 independent pre-treatment variables, with an accuracy of 69.4% and 79.3%, respectively: initial severity of anxiety and depression, ethnicity, deprivation and gender. The number of sessions attended/missed were also important factors identified in recovery.

Conclusions: Predicting whether patients are likely to have a positive outcome following treatment at entry might allow suitable modification of scheduled treatment, possibly resulting in improvements in outcomes. The model also highlights factors not only associated with poorer outcomes but inextricably linked to prevalence of common mental disorders, emphasising the importance of social determinants not only in poor health but also poor recovery.

Strengths and limitations of this study

- This article presents novel research exploring the effects of patient and treatment factors on the clinical outcomes of patients receiving psychological therapies as treatment for common mental disorders.
- The research has been developed with clinicians to contextualise and potentially apply findings to improve clinical outcomes of patients as an example of translational research.
- Assessing the likelihood of a patient to achieve a positive outcome at assessment may offer an opportunity to modulate/augment treatment or provide additional support to ensure the best possible outcome for the patient.
- While the sample size was sufficiently large to investigate the variables included in the study, the sample was not large enough to permit important subgroup analyses that could elucidate differences within ethnic groups, which are known to differ markedly in respect to common mental health problems, or possible interactions between different variables.
- Exclusion of patients who did not complete treatment was necessary to analyse treatment outcomes, but represents a significant proportion of those who entered the service. More needs to be done to understand the characteristics and underlying reasons for non-completion.

INTRODUCTION

The Improving Access to Psychological Therapies (IAPT) programme was launched in England in 2007 to provide community-based services for the treatment of people with common mental disorders (CMDs), including anxiety and depression.[1] The national programme supports the delivery of locally tailored psychological services, providing access to evidence-based psychological therapies, including those based on a

cognitive–behavioural therapy (CBT) approach, as recommended by national guidelines.[2 3] There is some evidence of increased prevalence of CMDs in certain populations including those living in areas of high socioeconomic deprivation, some members of Black and Minority Ethnic (BME) communities, forced migrants and asylum seekers.[4–6] Inequalities seen in these populations are often exacerbated by inequities in access to appropriate services, highlighting the need for specific strategies to improve access for these groups.[6–8]

The IAPT programme established a minimum data set for the routine collection of data including demographics, such as age, ethnicity, gender and residential postcode; information from patient interactions with the service, including treatment type/intensity and sessions attended or missed.[9] Clinical outcomes were assessed using the Patient Health Questionnaire-9 (PHQ-9) and Generalised Anxiety Disorder-7 (GAD-7) tools, administered before and after treatment.[10 11] The values of PHQ-9 and GAD-7 at entry determine a patient's 'caseness' or severity of depression or anxiety, respectively, which is the stated but not mandatory level for entry to IAPT (PHQ-9 ≥10 and GAD-7 ≥8). These measures are subsequently used to assess a positive response (recovery) following treatment, and is defined as those that have a PHQ-9 <10 and GAD-7 <8.[12]

These data have allowed periodical analyses and evaluations of the delivery of IAPT, including a report in 2012 that analysed data on the first 1 000 000 people referred to the service nationally.[13] The report concluded that the IAPT programme had provided people with access to evidence-based psychological therapies that they would not have previously accessed, and recovery rates of those treated (45%) were of similar magnitude to those seen in randomised controlled trials (50%), much of which could only be demonstrated through the complete and rigorous collection of key outcome measures.

Future developments of the programme include expansion of current services to specifically address the needs of those with long-term conditions and medically unexplained symptoms, and the development of new services to support children and young people, and those with a severe and enduring mental illness.[13]

Despite the apparent success of the programme, there is heterogeneity in treatment outcomes, with little evidence of what works for whom. A better understanding of characteristics associated with specific outcomes and the ability to predict the likelihood of patients achieving a positive outcome may offer an opportunity to provide more individualised treatment, and improve outcomes for all those who access the service.

This paper aims to identify patient and treatment factors that affect clinical outcome of community psychological therapy through the development of a predictive model using historic data from two IAPT services in London. In addition, the paper aims to assess the completeness of data collection, explore how treatment outcomes are discriminated using current criteria for classifying recovery using the predictive model, and assess the feasibility and need for undertaking a future larger population analysis.

METHOD

Ethical approval

Ethics approval was not required for this work as the clinical team providing care for patients anonymised all data routinely collected for clinical purposes, prior to transferring the data to researchers for analysis (in accordance with UK Governance Arrangements for Research Ethics Committees, Section C4). The research study was registered with Chelsea and Westminster NHS Foundation Trust.

Study setting

The anonymised data was extracted from two community-based IAPT services that serve two different London boroughs. Service A, launched in 2009, included a team of 30 therapists, serving a population of more than 200 000 people, where more than 30% of the local population is from a BME background, and under 60% of the working age population is employed. Service B, established in 2010, comprised a team of 15 therapists serving a population of nearly 300 000 people, of which up to 60% come from a BME background, and over 60% of the working age population is employed. Both boroughs include significant pockets of severe deprivation within the top 5% in UK.

Data sources/sample and exclusion criteria

Data for all referrals to the services between February 2009 and May 2012 were collected by the IAPT clinical team using IAPTus (Mayden, Wiltshire, UK), a clinical data system. Data included independent and dependent variables, as described below, collected as part of the minimum dataset, collected at two time points, during the first session (pre-treatment) and at the final session (post-treatment).[9]

To maintain anonymity, the clinical team converted postcode to *Index of Multiple Deprivation (IMD)* prior to transferring the data to researchers. Only cases with values for both final scores of *PHQ-9* and *GAD-7* were included in the analyses, as both were required for generating the outcome measures.

Patients assigned inappropriate values for *age* or *length of treatment* (ie, *age* <0 or >105; *length of treatment* >14 000 days—a value created by the database to indicate those where no start was indicated, as no length of treatment could be calculated) were removed from analysis. Data were imported into SPSS (IBM SPSS Statistics V.21), and where necessary, variables were converted from alphanumeric to numeric or coded data.

Independent variables

Independent variables were selected on the basis of availability within the data set and were classified

according to the temporal collection of data, that is, first session (pre-treatment) or final session (post-treatment). Pre-treatment included: *gender [Male/Female]*, *age* [1–105], *ethnicity [White-British/BME/Not stated]*, *able to communicate in spoken English* [Yes/No], *understand written English* [Yes/No], *source of referral* [GP/self-referral], *PHQ-9 first score* [0–27] and *GAD-7 first score* [0–21] and postcode converted to *Index of Multiple Deprivation (IMD)* [1–70]. The values for the *PHQ-9 first score* and *GAD-7 first score* were checked for 'caseness' that is, PHQ-9 ≥10 and GAD-7 ≥8. Post-treatment variables included: *number of sessions not attended (DNA)*, *number of sessions attended*, *length of treatment* [Days], *reason for end of IAPT care pathway* [planned ending/deceased/declined further contact/dropped out/ineligible for service/sign-posted/no treatment indicated], *Guided Self-Help* [Yes/No] and *high-intensity treatment* (Cognitive Behaviour Therapy) [Yes/No].

Outcome measures

Severity of anxiety and depression was assessed during the final treatment session using the GAD-7 tool and the PHQ-9 tool, respectively. The scores generated represented the dependent post-treatment variables *GAD-7 final score* and *PHQ-9 final score*.

These variables were used to allocate patients to a new categorical outcome variable, *Treatment Outcome*. *GAD-7 final score* less than eight and *PHQ-9 final score* less than 10 indicate positive outcomes, while scores ≥8 or ≥10, respectively, indicate negative outcomes.

This resulted in four treatment outcome options, P1P2, P1N2, N1P2 and N1N2, depending on whether either or both outcomes were positive (P) or negative (N), for example, those achieving a PHQ-9 <10 (P1) and GAD-7 <8 (P2) were allocated to the P1P2 Treatment Outcome, as shown in table 1.

Outcome group

Further to establishing an outcome measure for the study, these were further classified to create outcome groups, to allow separate analysis of the current approaches to classifying recovery. The analysis was based on Treatment Outcomes, testing P1P2 against N1N2 (Outcome Group 1), P1P2 against P1N2, N1P2 and N1N2 combined (Outcome Group 2) and P1P2 versus P1N2 versus N1P2 versus N1N2 (Outcome Group 3). The data cleaning and allocation to outcome groups is outlined in figure 1.

Table 1 Classification of treatment outcome combining final values of PHQ-9 and GAD-7

Treatment outcome	Final PHQ-9		Final GAD-7	
	<10	≥10	<8	≥8
P1P2	Positive		Positive	
P1N2	Positive			Negative
N1P2		Negative	Positive	
N1N2		Negative		Negative

GAD-7, Generalised Anxiety Disorder-7; N1P2, negative PHQ-9, positive GAD-7; N1N2, negative PHQ-9, negative GAD-7; PHQ-9, Patient Health Questionnaire-9; P1P2, positive PHQ-9, positive GAD-7; P1N2, positive PHQ-9, negative GAD-7.

Assessing completeness of data and a comparison of data sets

Frequencies were calculated for the categorical data and descriptive statistics (mean, SD, minimum and maximum) for the numerical variables from each service. Descriptive statistics for the numerical variables were calculated for the combined data, classified by Treatment Outcome: P1P2, P1N2, N1P2 and N1N2. The non-parametric independent samples Kruskal-Wallis test (K-W test), and the univariate Analysis of Variance (univariate ANOVA) procedure were used to check for any differences in the variables: *Age*, *Number of sessions not attended (DNA)*, *Number of sessions attended*, *Length of Treatment*, *PHQ9 first score*, *GAD7 first score* and IMD, comparing the four Treatment Outcomes. The post hoc test, Tukey's honestly significant difference was used to identify any significant differences between the outcome groups.[14]

Identifying predictors and developing a predictive model

The statistical procedure used for both the prediction of treatment outcome and also to identify which variables contributed to a positive treatment outcome was *Classify by means of Discriminant Analysis*.[15 16] This procedure builds a predictive model for group membership. The model is composed of a discriminant function (or, for more than two groups, a set of discriminant functions) based on linear combinations of the predictor variables that provide the best discrimination between the groups. For the initial calculations, the functions were generated from a sample of randomly selected cases for which group membership was known; the functions could then be tested on the unselected cases with known group membership. If the functions produced correct predictions for 60% or above, the model was accepted as suitable for the predicting cases with unknown group membership.

The grouping variable can have more than two values. The codes for the grouping variable must be integers, and it was necessary to specify the respective minimum and maximum values. Patients with values outside of these bounds were excluded from the analysis. The tests of equality of group means, Wilks' λ and the significance test, were used to identify the relative contribution of each variable to the models; the lower the value of Wilks' λ, the greater the contribution to the model.

Discriminant Analysis using only the pre-treatment variables known was used in order to test the possibility of assessing whether a patient would have a positive or negative response to the standard treatment schedule. The same statistical procedure, using variables known at the completion of treatment (pre-treatment and post-treatment variable), was used to identify which variables

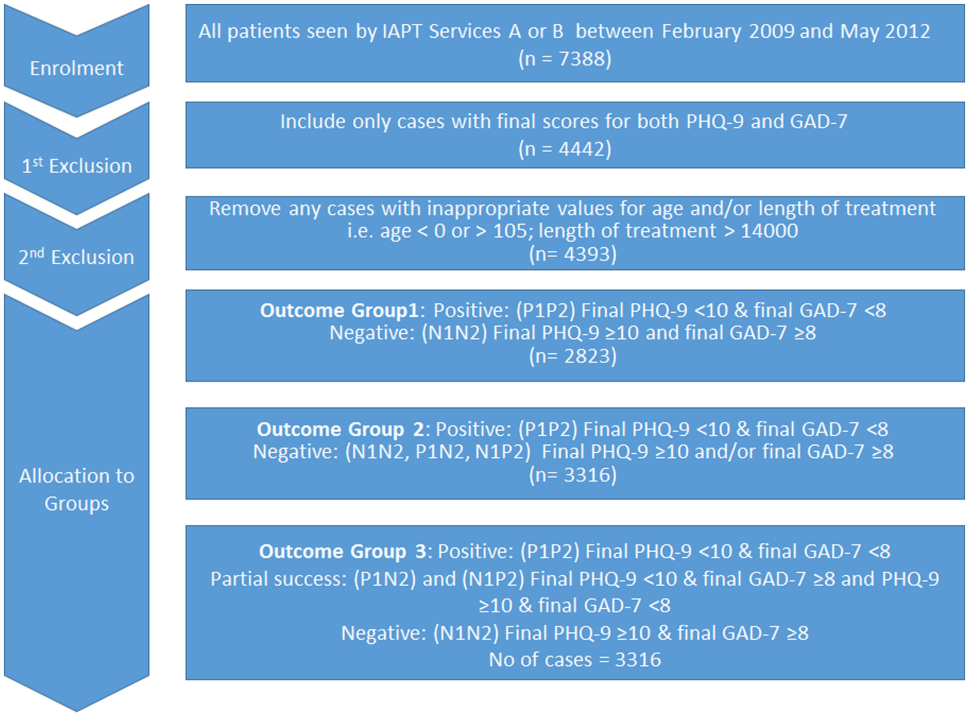

Figure 1 Flow diagram showing the organisation of procedures and the number of patients included at each stage of data cleaning and allocation of patients to outcome groups (GAD-7, Generalised Anxiety Disorder-7; IAPT, Improving Access to Psychological Therapies; N1P2, negative PHQ-9, positive GAD-7; N1N2, negative PHQ-9, negative GAD-7; PHQ-9, Patient Health Questionnaire-9; P1P2, positive PHQ-9, positive GAD-7; P1N2, positive PHQ-9, negative GAD-7).

contributed to the models, and were therefore influencing whether a patient had a positive or negative Treatment Outcome.

In order to test whether the model produced by Discriminant Analysis was consistent for different services, Services A and B were each used to classify both services. The results were then compared to check the percentage of cases with identical classifications.

RESULTS

Assessing completeness of data and a comparison of data sets

The frequencies for the categorical variables for Services A and B are shown in table 2. The analysis demonstrated some interesting results especially relating to the proportions of patients with a BME background attending the services, (A=50%, B=63%) compared with white British

Table 2 Frequencies of categorical variables comparing services A and B

Variable	Completeness of data recording (%)	Service A N (%)	Service B N (%)
Gender M:F	100	1098:2061 (35:65)	445:787 (36:64)
Ethnicity White British: BME: not stated	93	1292:1438:131 (45:50:5)	361:776:95 (29:63:8)
Able to communicate in spoken English? Yes:No:Don't know	91	2828:39:32 (98.3:0.9:0.9)	1048:39:18 (94.8:3.5:1.6)
Understand written English? Yes:No:Don't know	90	2828:39:32 (97.6:1.3:1.1)	972:43:43 (91.9:4.1:4.1)
Source of Referral GP:Self:Other	100	2268:656:236 (71.8:20.8:7.5)	1157:2:73 (93.9:0.2:5.9)
Caseness threshold met Yes:No	100	2130:1029 (67.4:32.6)	885:347 (71.8:28.2)
GSH Yes:No	100	2110:1050 (66.8:33.2)	455:777 (36.9:63.1)
CBT Y:N	100	1651:1509 (52.2:47.8)	822:410 (66.7:33.3)

BME, black and minority ethnic; CBT, cognitive–behavioural therapy; GSH, guided self-help.

patients (A=45%, B=29%). Patients attending Service B were less likely to have entered the GSH programme (A=66.8%, B=36.9%) and more likely to have had CBT (A=52.2%, B=66.7%) than patients from Service A. More patients attending Service A were self-referred (A=20.8% B=0.2%), while the majority of patients attending Service B were GP referred (A=71.8%, B=93.9%). There did not appear to be any difference between the two Services with regard to the ratio of males to females, ability to communicate in spoken English, and to understand written English. The percentage of patients assessed to be 'case' at admission was similar in both Services (A=67.4%, B=71.8%).

The descriptive statistics for the numeric variables for Services A and B are shown in table 3. Apart from the age range of the patients, and first and last PHQ-9/GAD-7 scores, which were similar for both Services, the values of variables for Service B tended to be numerically lower than those for Service A, although the ranges were similar.

The descriptive statistics for the numeric variables by Treatment Outcome are shown in table 4. Although the age ranges are similar, there is a slight increase in the mean ages from the positive to the negative Treatment Outcome. The entry scores for PHQ-9 and GAD-7 are both lower for the patients with a positive, compared with negative Treatment Outcome.

Identifying predictors and developing a predictive model

Are there sufficient differences between the treatment outcomes?

The K-W test showed that the distribution of variables known at initial enrolment including PHQ-9/GAD-7 scores, Age and IMD, were not the same across the categories of Treatment Outcome. The distribution of post-treatment variables including length of treatment, number of sessions attended and DNA were also shown to be different. OneWay ANOVA and the Tukey's HND test were used to identify the differences within the pre-treatment and post-treatment variables, as a substitute for the Student-Newman-Keuls procedure (table 5).[17] These results suggested that there were sufficient differences between the positive and negative Treatment Outcomes to enable the Discriminant Analysis to differentiate between them. The table also demonstrates the contribution of numerical direction of variables that is, high values for PHQ-9/GAD-7 first scores, IMD, are associated with poorer outcomes (4th column, highest value shows N1N2), while a high value for number of sessions attended indicates a positive outcome (4th column highest value shows P1P2).

Outcome group 1: can positive and negative treatment outcomes be predicted from pre-treatment variables?

Discriminant Analyses for Outcome Group 1 data (differentiate between P1P2 and N1N2) using nine pre-treatment variables: *gender*, *age*, *ethnicity*, *caseness achieved*, *understand written English*, *source of referral*, *PHQ-9 first score*, *GAD-7 first score* and IMD showed that the Treatment Outcome (P1P2, N1N2) could be correctly predicted for between 69.9–76% of cases, table 6. Classification of the selected grouped cases (ie, cases with known Treatment Outcome used to calculate the model) was correct for 71.8%, and for the unselected original grouped cases (ie, cases with known Treatment Outcome not used to calculate, but to test the model) was 74.9% correct. The classifications of the selected

Table 3 Descriptive statistics for numeric variables for services A and B

	Completeness of data recording (%)	Service	n	Minimum	Maximum	Mean	SD
Age	100	A	3159	17	100	38.3	13.3
		B	1232	12	95	38.2	12.7
Length of treatment (days)	100	A	3160	0	571	104.5	86.1
		B	1232	0	409	96.6	79.0
DNA sessions (n)	100	A	3160	0	12	0.71	1.2
		B	1232	0	9	0.8	1.2
Attended sessions (n)	100	A	3160	1	33	7.5	5.4
		B	1232	1	30	6.5	4.6
PHQ9 first score	100	A	3160	0	27	14.3	6.5
		B	1232	0	27	15.4	6.8
PHQ9 last score	100	A	3160	0	27	9.2	6.9
		B	1232	0	27	10.9	7.4
GAD-7 first score	100	A	3158	0	21	12.6	5.6
		B	1232	0	21	13.4	5.5
GAD-7 last score	100	A	3160	0	21	8.3	6.0
		B	1232	0	21	9.7	6.2
IMD	98.2	A	3095	1.4	69.7	27.1	12.6
		B	1221	9.7	61.4	29.9	11.2

DNA, sessions not attended; GAD-7, Generalised Anxiety Disorder-7; IMD, Index of Multiple Deprivation; PHQ9, Patient Health Questionnaire-9.

Table 4 Descriptive statistics for numeric variables for each treatment outcome

Pre-treatment/post-treatment variable	Variable	Treatment outcome	n	Minimum	Maximum	Mean	SD
Pre-treatment	Age	P1P2	2001	17	86	37.86	13.36
		P1N2	440	18	88	36.67	13.42
		N1P2	222	17	83	40.17	14.97
		N1N2	1728	12	100	38.92	12.51
Pre-treatment	PHQ9 1st score	P1P2	2001	0	27	11.7	6.3
		P1N2	440	1	27	13.0	5.8
		N1P2	222	0	27	15.9	5.3
		N1N2	1729	0	27	18.3	5.3
Pre-treatment	GAD-7 1st score	P1P2	2000	0	21	10.6	5.5
		P1N2	439	0	21	13.6	4.7
		N1P2	222	0	21	11.1	5.4
		N1N2	1729	0	21	15.6	4.4
Pre-treatment	IMD	P1P2	1971	1.4	61.4	26.39	11.72
		P1N2	426	2.4	61.4	26.69	12.04
		N1P2	217	4.0	61.4	27.38	12.79
		N1N2	1702	3.1	69.7	29.94	12.66
Post-treatment	Length of treatment (days)	P1P2	1825	3	571	121.64	78.15
		P1N2	371	5	510	130.18	91.35
		N1P2	196	1	454	108.57	80.67
		N1N2	1448	1	482	108.94	79.06
Post-treatment	Number of sessions not attended	P1P2	2001	0	9	0.55	0.98
		P1N2	440	0	7	0.66	1.07
		N1P2	222	0	6	0.85	1.11
		N1N2	1729	0	12	0.97	1.31
Post-treatment	Number of sessions attended	P1P2	2001	1	32	8.01	5.13
		P1N2	440	1	32	7.14	5.54
		N1P2	222	1	28	6.19	4.76
		N1N2	1729	1	33	6.41	5.18

GAD-7, Generalised Anxiety Disorder-7; IMD, Index of Multiple Deprivation; N1P2, negative PHQ-9, positive GAD-7; N1N2, negative PHQ-9, negative GAD-7; P1P2, positive PHQ-9, positive GAD-7; P1N2, positive PHQ-9, negative GAD-7; PHQ9, Patient Health Questionnaire-9.

cases were also tested by cross-validation, where each case was classified by the functions derived from all cases other than that case, which showed that 71.6% of these classifications were correct.

The contribution of each variable to the model is shown in table 7; all the variables entered contributed to the model, but *PHQ-9 first score* and *GAD-7 first score* were the most important, followed by *caseness achieved*, IMD and *ethnicity*.

Outcome group 2: can positive and negative (including partial) treatment outcomes be predicted from pre-treatment variables?

Discriminant Analyses for Outcome Group 2 data (differentiate between P1P2 and all negative Treatment Outcomes, P1N2, N1P2 and N1N2 merged) using nine pre-treatment variables: *gender*, *age*, *ethnicity*, *caseness achieved*, *understand written English*, *source of referral*, *PHQ-9 first score*, *GAD-7 first score* and IMD showed that

Table 5 Grading from lowest to highest numerical levels for pre-treatment and post-treatment variables

Pre-treatment/post-treatment variable	Variable	Grading in order from lowest to highest level			
		1st	2nd	3rd	4th
Pre-treatment	Age	P1N2	P1P2	N1N2	N1P2
Pre-treatment	PHQ-9 1st score	P1P2	P1N2	N1P2	N1N2
Pre-treatment	GAD-7 1st score	P1P2	N1P2	P1N2	N1N2
Pre-treatment	IMD	P1P2	P1N2	N1P2	N1N2
Post-treatment	LoT (days)	N1P2	N1N2	P1P2	P1N2
Post-treatment	Number of DNA sessions	P1P2	P1N2	N1P2	N1N2
Post-treatment	Number of Attended sessions	N1P2	N1N2	P1N2	P1P2

GAD7, Generalised Anxiety Disorder-7; IMD, Index of Multiple Deprivation; LoT, length of treatment; N1P2, negative PHQ-9, positive GAD-7; N1N2, negative PHQ-9, negative GAD-7; P1P2, positive PHQ-9, positive GAD-7; P1N2, positive PHQ-9, negative GAD-7; PHQ9, Patient Health Questionnaire-9.

Table 6 Outcome Group 1 classification results for predictive model created using pre-treatment variables

Classification results				
	Treatment outcome P1P2 vs N1N2	Predicted group membership		
		P1P2 (%)	N1N2 (%)	Total (%)
Cases selected				
Original (71.8%)	P1P2	**69.9**	30.1	100.0
	N1N2	25.9	**74.1**	100.0
Cross-validated (71.6%)	P1P2	**69.9**	30.1	100.0
	N1N2	26.2	**73.8**	100.0
Cases not selected				
Original (74.9%)	P1P2	**73.9**	26.1	100.0
	N1N2	24.0	**76.0**	100.0

GAD7, Generalised Anxiety Disorder-7; N1P2, negative PHQ-9, positive GAD-7; N1N2, negative PHQ-9, negative GAD-7; P1P2, positive PHQ-9, positive GAD-7; P1N2, positive PHQ-9, negative GAD-7; PHQ9, Patient Health Questionnaire-9.

Treatment Outcome (P1P2, Negative) could be correctly predicted for between 66.8% and 71.0% of cases, table 8. Classification of the selected grouped cases was correct for 68.5%, and for the unselected original grouped cases it was 70.9% correct. The classifications of the selected cases tested by cross-validation showed that 68.3% of these classifications were correct.

The contribution of each variable to the model is shown in table 9; seven of the variables entered contributed to the model, *caseness achieved* and *source of referral* were omitted. *PHQ-9 first score* and *GAD-7 first score* were the most important, followed by IMD and *ethnicity*.

Outcome group 2: can positive and negative (including partial) treatment outcomes be predicted from pre-treatment and post-treatment variables?

Discriminant Analyses for Outcome Group 2 data (differentiate between P1P2 and all negative Treatment Outcomes, P1N2, N1P2 and N1N2 merged) using 14 pre-treatment and post-treatment variables: *gender*, *age*, *ethnicity*, *understand written English*, *source of referral*, *PHQ-9 first score*, *GAD-7 first score*, IMD, *length of treatment*, *number of sessions not attended (DNA)*, *number of sessions attended*, *reason for end of IAPT care pathway*, *Guided Self-Help* and *high intensity treatment*, showed that Treatment Outcome (P1P2, Negative) could be correctly predicted for between 66.8% and 71.0% of cases, table 10. Classification of the selected grouped cases was correct for 70.7%, and for the unselected original grouped cases it was 71.3% correct. The classifications of the selected cases tested by cross-validation showed that 70.2% of these classifications were correct.

The contribution of each variable to the model is shown in table 11; 11 of the variables entered contributed to the model, while age, source of referral and reason for end of IAPT care pathway were omitted. The variables that had the most effect on the Treatment Outcome were, in order of importance, PHQ9 and GAD-7 1st scores, followed by DNA, number of attended sessions, IMD and ethnicity. The length of treatment, CBT, understanding of written English, Guided Self-Help and gender, although contributing, were of least effect.

Table 7 Contributions of individual pre-treatment variables to the model created to predict outcomes for the model for predicting outcomes within Outcome Group 1

Tests of equality of group means			
Order of Importance	Variables	Wilks' λ	Significance
1	PHQ-9 first	0.794	<0.001
2	GAD-7 first score	0.823	<0.001
3	Caseness achieved	0.849	<0.001
4	IMD	0.976	<0.001
5	Ethnicity	0.987	<0.001
6	Age	0.995	0.005
7	Source of referral (coded)	0.996	0.010
8	Gender	0.997	0.039
9	Understand written English?	0.998	0.058

GAD7, Generalised Anxiety Disorder-7; IMD, Index of Multiple Deprivation; PHQ9, Patient Health Questionnaire-9.

Outcome group 3: can positive, partial and negative treatment outcomes be predicted from pre-treatment variables?

Discriminant Analyses for Outcome Group 3 data (differentiate between P1P2, P1N2, N1P2 and N1N2) using the eight pre-treatment variables: *gender*, *age*, *ethnicity*, *understand written English*, *source of referral*, *PHQ-9 first score*, *GAD-7 first score* and IMD showed that Treatment Outcome (P1P2, P1N2, N1P2 and N1N2) could be correctly predicted for between 38.9% and 56.3% of cases, table 12. These results are no better than that obtained by chance and are not acceptable. Therefore, although it is possible to produce a model that will discriminate between positive and negative Treatment Outcomes, it is not possible to distinguish between the partial and total negative Treatment Outcomes using the variables available at enrolment.

Table 8 Outcome Group 2 classification results for predictive model created using pre-treatment variables

Classification results				
	Treatment outcome P1P2 vs 3 other groups merged	Predicted group membership P1P2 (%)	N and partial N (%)	Total (%)
Cases selected				
Original (68.5%)	P1P2	**67.2**	32.8	100.0
	P1N2, N1P2 and N1N2	30.3	**69.7**	100.0
Cross-validated (68.3%)	P1P2	**66.8**	33.2	100.0
	P1N2, N1P2 and N1N2	30.4	**69.6**	100.0
Cases not selected				
Original (70.9%)	P1P2	**71.0**	29.0	100.0
	P1N2, N1P2 and N1N2	29.3	**70.7**	100.0

GAD7, Generalised Anxiety Disorder-7; N1P2, negative PHQ-9, positive GAD-7; N1N2, negative PHQ-9, negative GAD-7; P1P2, positive PHQ-9, positive GAD-7; P1N2, positive PHQ-9, negative GAD-7; PHQ9, Patient Health Questionnaire-9.

Comparing the results when data from Service A or Service B was used to classify the Treatment Outcome, 2532 cases of the 3830 (66.1%) in the database gave identical results. This is acceptably close to the overall percentage of cases correctly classified, justifying merging the data from the two services. The model is therefore robust enough to allow calculations on data from one service to be made using a model based on another service.

DISCUSSION

The analysis identified initial severity of anxiety and depression, ethnicity, deprivation and gender as pre-treatment predictors of recovery. The study also demonstrates the importance of the duration of treatment as seen by the relationship between number of sessions attended/missed and recovery. The predictive models developed were able to assign a positive or negative clinical outcome to each patient based on these five independent pre-treatment variables, with an accuracy of 69.4% and 79.3%, respectively. The assessment of completeness of data collection has established the accuracy with which analysis of IAPT data can be undertaken, and the feasibility and necessity of undertaking larger scale analysis of population data to specifically assess the situation for recovery of those from BME groups, areas of deprivation and, importantly, further understand the reasons why a significant number of patients 'drop out' from the service.

Table 9 Contributions of individual pre-treatment variables to the model created to predict outcomes for the model for predicting outcomes within Outcome Group 2

Tests of equality of group means			
Order of importance	Variable	Wilks' λ	Significance
1	PHQ-9 first score	0.846	<0.001
2	GAD-7 first score	0.858	<0.001
3	IMD	0.986	<0.001
4	Ethnicity	0.992	<0.001
5	Understand written English?	0.998	0.036
6	Gender	0.999	0.045
7	Age	0.999	0.077

GAD7, Generalised Anxiety Disorder-7; IMD, Index of Multiple Deprivation; PHQ9, Patient Health Questionnaire-9.

All IAPT services in England collect, collate and analyse patient-level data to provide individualised feedback to patients on progress and monitor service performance. Despite the availability of this rich data set, very little work has been done to explore the factors associated with why some people recover following treatment within the IAPT service and others do not. It is likely to become increasingly important for individual IAPT services to understand their own patient and service/treatment characteristics associated with enhanced recovery rates, to ensure continued outcomes, funding and support, as targets for recovery increase. Understanding how the potential differences in local population composition could impact on outcomes will also help understand differences on key performance indicators between IAPT services.

Gyani *et al*[18] published a report summarising recovery rates and their predictors across 32 IAPT sites during the first year of their operation. A multivariate logistic regression found that a number of determinants were significantly associated with the likelihood of recovery across sites. First, greater numbers of therapy sessions were associated with higher recovery rates. Second, severity of initial symptoms had a negative impact on likelihood of recovery; the higher the initial PHQ-9 and GAD-7 scores, the less likely a patient was to achieve recovery. Third, experience/seniority of therapists had an impact on likelihood of recovery, with higher levels of experience/seniority associated with more successful outcomes. While Gyani *et al*[18] provide some insights echoed by the current analysis, patient-level characteristics such as age, sex, ethnicity and language ability, were not included in their analysis of predictors of outcome.

Table 10 Outcome Group 2 classification results for predictive model created using pre-treatment and post-treatment variables

Classification results				
	Treatment outcome P1P2 vs 3 other groups merged	Predicted group membership P1P2 (%)	N and partial N (%)	Total
Cases selected				
Original (70.7%)	P1P2	**71.4**	28.6	100.0
	P1N2, N1P2 and N1N2	30.0	**70.0**	100.0
Cross-validated (70.2%)	P1P2	70.9	29.1	100.0
	P1N2, N1P2 and N1N2	30.4	**69.6**	100.0
Cases not selected				
Original (71.3%)	P1P2	**69.3**	30.7	100.0
	P1N2, N1P2 and N1N2	26.8	**73.2**	100.0

GAD7, Generalised Anxiety Disorder-7; N1N2, negative PHQ-9, negative GAD-7; N1P2, negative PHQ-9, positive GAD-7; P1P2, positive PHQ-9, positive GAD-7; P1N2, positive PHQ-9, negative GAD-7; PHQ9, Patient Health Questionnaire-9.

The findings from this study, similar to that of Gyani *et al*[18], show that the more sessions a patient received, the more likely they were to recover. Also, that initial symptom severity had a negative impact on likelihood of recovery, the higher the initial PHQ-9 and GAD-7 scores, the less likely a patient was to achieve recovery. Gyani *et al*[18] suggests that this is an artefact of the way that recovery is defined, as, in fact, IAPT services can offer benefits to patients across a spectrum of severity.[18]

Previous studies have also reported that while psychological therapies can demonstrate an improvement in clinical outcomes, more than eight sessions achieve a greater positive response, in those with anxiety at least.[19]

Table 11 Contributions of individual pre-treatment and post-treatment variables to the model created to predict outcomes for the model for predicting outcomes within Outcome Group 1

Tests of equality of group means			
Order of importance	Variables	Wilks' λ	Significance
1	PHQ9 1st score	0.842	<0.001
2	GAD7 1st score	0.865	<0.001
3	No. DNA sessions	0.974	<0.001
4	No. attended sessions	0.981	<0.001
5	IMD	0.987	<0.001
6	Ethnicity	0.992	<0.001
7	Length of treatment in days	0.998	0.041
8	Cognitive–behavioural therapy	0.998	0.037
9	Understand written English?	0.998	0.057
10	Guided self help	0.999	0.077
11	Gender	0.999	0.087

GAD7, Generalised Anxiety Disorder-7; IMD, Index of Multiple Deprivation; PHQ9, Patient Health Questionnaire-9.

Furthermore, this study has demonstrated that it is possible to develop a robust model based on discriminant analysis that is accurate in predicting the outcome of patients given the specific independent variables prior to the start of treatment. While the model is able to successfully discriminate those patients with a positive outcome (P1P2) from those with a negative outcome (N1N2), it was less accurate in distinguishing those patients who achieved a positive outcome in either PHQ-9 or GAD-7 scores alone (P1N2 or N1P2). Factors identified that were associated with outcome were severity of disease on entry (initial PHQ-9 and GAD-7 scores), ethnicity, socioeconomic status (IMD) and gender. Analysis of data has demonstrated a high level of completion with an average of 93.2% completion rate for the data fields.

The effect of factors including ethnicity and gender on clinical outcomes following treatment of CMD with psychological therapies in the literature is scant or conflicting. There seems to be a dearth of studies that have specifically aimed to assess the effect of these demographic factors on outcomes.[20] Our group has previously demonstrated an association between socioeconomic status and severity of illness; while that study demonstrated the ability for patients from deprived areas to achieve similar outcomes as those from less deprived areas, this was in response to a programme of activity targeted at this population.[21] Socioeconomic status, as assessed by educational attainment and income, was also demonstrated to be associated with greater improvements in clinical outcomes.[22]

Referral rates in the study are similar to those reported elsewhere with females accounting for 61%, and those classed as white British accounting for 37.6%.[18 23] The well described link between prevalence in CMDs and deprivation is noted in the reduced likelihood of a positive treatment outcome in patients from areas of higher deprivation.[4 24] While deprivation has been shown to be associated with outcomes, the effect may well be underestimated.

Table 12 Outcome Group 3 classification results for predictive model created using pre-treatment variables

Classification results						
	Treatment outcome 4 choices	Predicted group membership				
		P1P2	P1N2	N1P2	N1N2	Total
Cases selected						
Original	P1P2	**40.4**	23.7	20.4	15.5	100.0
	P1N2	22.5	**42.6**	9.3	25.6	100.0
	N1P2	19.1	9.2	**40.5**	31.3	100.0
	N1N2	9.5	19.8	18.5	**52.2**	100.0
Cross-validated	P1P2	**40.2**	23.8	20.4	15.5	100.0
	P1N2	22.5	**42.2**	9.3	26.0	100.0
	N1P2	20.6	9.2	**38.9**	31.3	100.0
	N1N2	9.5	20.1	18.7	**51.7**	100.0
Cases not selected						
Original	P1P2	**40.8**	24.4	19.8	15.0	100.0
	P1N2	21.9	**47.9**	11.5	18.8	100.0
	N1P2	17.2	6.3	**56.3**	20.3	100.0
	N1N2	11.4	18.3	19.6	**50.7**	100.0

GAD7, Generalised Anxiety Disorder-7; N1P2, negative PHQ-9, positive GAD-7; N1N2, negative PHQ-9, negative GAD-7; P1P2, positive PHQ-9, positive GAD-7; P1N2, positive PHQ-9, negative GAD-7; PHQ9, Patient Health Questionnaire-9.

Attrition has previously been shown to be linked to deprivation suggesting that patients from areas of higher deprivation are more likely to 'drop out' of the service, and thus potentially excluded from the analysis.[25]

Our previous study of a subset of the data demonstrated that patient outcomes were similar across all deprivation groups, although that analysis used PHQ-9 as the sole outcome measure.[21]

Limitations of this study

A proportion of the patients were excluded from the study due to incomplete data for certain independent variables of interest. Imputation was not considered appropriate to recover these exclusions, as we had no specific data to provide 'average' values. Additionally, while the sample size was sufficiently large to investigate the variables included in the study, a larger study would allow subanalysis of variables and their interactions with each other including factors such as different ethnicity categories and intensity/mode of treatment. Importantly, a significant population of patients who were referred to the service have been excluded from the analysis, those who do not complete treatment, that is, those who 'dropped out'. It is clear that it is extremely important to understand the reasons why this group does not complete treatment, although this methodology may have limitations in accurately characterising this population. The authors are currently undertaking a larger study to ensure this important subanalysis (including those who drop out) is undertaken in collaboration with a larger number of services across the northwest London area.

CONCLUSION

As De Lusignan *et al*[23] has demonstrated, people with CMDs have a higher rate of healthcare resource utilisation than those without CMDs, and IAPT referral can actively reduce this along with a reduction in sickness reporting, as was intended by the programme. While previous studies have demonstrated a range of interventions that can be harnessed to increase access to IAPT services, more needs to be done to tailor treatment to improve the outcomes of individuals.[21 26 27] Interventions and treatment moderators should be identified that improve support and engagement for those 'at risk' of not achieving a positive outcome following treatment. Identifying patients at risk of achieving poor outcomes at entry offers an opportunity to provide enhanced support for this group, which might include the further development of culturally sensitive services or additional support relevant to those living in deprived areas, including better access through improvements in transport or patient incentives to encourage attendance.

Further studies should be developed to specifically investigate the exact nature and extent to which deprivation influences clinical outcomes, and potential interventions developed to ameliorate any negative effect. While there is some evidence of an increased prevalence of CMDs in some BME groups, a more detailed analysis would be required to dissect out the individual ethnic groups from 'BME' and cross-link this data with information on deprivation or communication skills to ensure a more sophisticated characterisation of the population, not just a typology based on one characteristic such as ethnicity.

Twitter Follow Stuart Green at @stuart4clahrc

Contributors SAG, JG and GP conceived the study. EH provided the data. SAC, AJP and TW contributed to the planning. SAC undertook the analysis. SAG and EH wrote the first draft and all other authors contributed to the final draft.

Funding This work was supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) Northwest London and the Central and North West London NHS Foundation Trust.

Competing interests This article presents independent research commissioned by the National Institute for Health Research (NIHR) under the Collaborations for Leadership in Applied Health Research and Care (CLAHRC) programme North West London.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1. CSIP Choice and Access Team. *IAPT positive practice guide*. London, 2007.
2. National Institute for Health and Care Excellence. *Depression in adults: the treatment and management of depression in adults*. London, 2009.
3. National Institute for Health and Care Excellence. *Generalised anxiety disorder and panic disorder (with or without agoraphobia) in adults: management in primary, secondary and community care*. London, 2011.
4. Fryers T, Melzer D, Jenkins R. Social inequalities and the common mental disorders: a systematic review of the evidence. *Soc Psychiatry Psychiatr Epidemiol* 2003;38:229–37.
5. Weich S, Nazroo J, Sproston K, *et al*. Common mental disorders and ethnicity in England: the EMPIRIC study. *Psychol Med* 2004;34:1543–51.
6. Palmer D, Ward K. 'Lost': listening to the voices and mental health needs of forced migrants in London. *Med Confl Surviv* 2007;23:198–212.
7. Commander MJ, Dharan SP, Odell SM, *et al*. Access to mental health care in an inner-city health district. II: Association with demographic factors. *Br J Psychiatry* 1997;170:317–20.
8. Kovandžić M, Chew-Graham C, Reeve J, *et al*. Access to primary mental health care for hard-to-reach groups: from 'silent suffering' to 'making it work'. *Soc Sci Med* 2011;72:763–72.
9. National IAPT Programme Team. *The IAPT data handbook. Appendices*. London. 2011.
10. Kroenke K, Spitzer RL. The PHQ-9: a new depression measure. *Psychiatr Ann* 2002;32:509–15.
11. Spitzer RL, Kroenke K, Williams JBW, *et al*. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006;166:1092–7.
12. National IAPT Programme Team. *IAPT Key Performance Indicator (KPI) technical guidance for adult IAPT services*. London: 2012.
13. Department of Health. *IAPT three-year report the first million patients*. London: 2012.
14. Seaman MA, Levin JR, Serlin RC. New developments in pairwise multiple comparisons: Some powerful and practicable procedures. *Psychol Bull* 1991;110:577–86.
15. Huberty CJ. *Applied Discriminant Analysis*. 1st edn. John Wiley & Sons, 1994.
16. Norusis M. *SPSS 13.0 Statistical Procedures Companion*. Upper Saddle River, NJ: Prentice Hall, Inc, 2004.
17. Keuls M. The use of the 'studentized range' in connection with an analysis of variance. *Euphytica* 1952;1:112–22.
18. Gyani A, Shafran R, Layard R, *et al*. Enhancing Recovery Rates in IAPT Services: Lessons from analysis of the Year One data. 2011.
19. Hunot V, Churchill R. Psychological therapies for generalised anxiety disorder. *Cochrane Database Syst Rev* 2007;(1):CD001848.
20. Parker G, Blanch B, Crawford J. Does gender influence response to differing psychotherapies by those with unipolar depression? *J Affect Disord* 2011;130:17–20.
21. Poots AJ, Green SA, Honeybourne E, *et al*. Improving mental health outcomes: achieving equity through quality improvement. *Int J Qual Health Care* 2014;26:198–204.
22. Pirkis J, Bassilios B, Fletcher J, *et al*. Clinical improvement after treatment provided through the Better Outcomes in Mental Health Care (BOiMHC) programme: do some patients show greater improvement than others? *Aust N Z J Psychiatry* 2011;45:289–98.
23. De Lusignan S, Chan T, Parry G, *et al*. Referral to a new psychological therapy service is associated with reduced utilisation of healthcare and sickness absence by people with common mental health problems: a before and after comparison. *J Epidemiol Community Health* 2012;66:e10.
24. Fryers T, Melzer D, Jenkins R, *et al*. The distribution of the common mental disorders: social inequalities in Europe. *Clin Pract Epidemiol Ment Health* 2005;1:14.
25. Self R, Oates P, Pinnock-Hamilton T, *et al*. The relationship between social deprivation and unilateral termination (attrition) from psychotherapy at various stages of the health care pathway. *Psychol Psychother* 2005;78:95–111.
26. Green SA, Poots AJ, Marcano-Belisario J, *et al*. Mapping mental health service access: achieving equity through quality improvement. *J Public Health (Oxf)* 2013;35:286–92.
27. Evans L, Green S, Sharma K, *et al*. Improving access to primary mental health services: are link workers the answer? *London J Prim Care (Abingdon)* 2014;6:23–8.