ISSN 2042-2695

CENTRE for ECONOMIC PERFORMANCE

CEP Discussion Paper No 1227

July 2013

Enhancing Recovery Rates: Lessons from Year One of the English Improving Access to Psychological Therapies Programme

Alex Gyani
Roz Shafran
Richard Layard
David M Clark

LSE THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Abstract

Background: The English Improving Access to Psychological Therapies (IAPT) initiative aims to make evidence-based psychological therapies for depression and anxiety disorder more widely available in the National Health Service (NHS). 32 IAPT services based on a stepped care model were established in the first year of the programme. We report on the reliable recovery rates achieved by patients treated in the services and identify predictors of recovery at patient level, service level, and as a function of compliance with National Institute of Health and Care Excellence (NICE) Treatment Guidelines.

Method: Data from 19,395 patients who were clinical cases at intake, attended at least two sessions, had at least two outcomes scores and had completed their treatment during the period were analysed. Outcome was assessed with the patient health questionnaire depression scale (PHQ-9) and the anxiety scale (GAD-7).

Results: Data completeness was high for a routine cohort study. Over 91% of treated patients had paired (pre-post) outcome scores. Overall, 40.3 % of patients were reliably recovered at post-treatment, 63.7% showed reliable improvement and 6.6% showed reliable deterioration. Most patients received treatments that were recommended by NICE. When a treatment not recommended by NICE was provided, recovery rates were reduced. Service characteristics that predicted higher reliable recovery rates were: high average number of therapy sessions; higher step-up rates among individuals who started with low intensity treatment; larger services; and a larger proportion of experienced staff.

Conclusions: Compliance with the IAPT clinical model is associated with enhanced rates of reliable recovery.

Keywords: Mental illness, depression, anxiety, cognitive behavioural therapy
JEL Classifications: I12, I18

This paper was produced as part of the Centre's Wellbeing Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

Acknowledgements

Alex Gyani is a Research Student at the School of Psychology and Clinical Language Sciences, University of Reading. Roz Shafran is a Professor at the School of Psychology and Clinical Language Sciences, University of Reading. Richard Layard is Director of the Wellbeing Programme at the Centre for Economic Performance and Emeritus Professor of Economics, London School of Economics and Political Science. David M Clark is a Professorial Fellow with the Department of Experimental Psychology, University of Oxford.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

Introduction

There is substantial evidence that cognitive behaviour therapy (CBT) and several other psychological therapies are effective treatments for depression and/or anxiety disorders. Starting in 2004 the National Institute of Health and Care Excellence (NICE) conducted systematic reviews of research on the efficacy of interventions for depression and anxiety disorders. The reviews led to the publication of a series of clinical guidelines that advocate the use of specific forms of CBT for depression and all the anxiety disorders (NICE 2004a, 2004b, 2005a, 2005b, 2006, 2009a, 2009b, 2011, 2013.) Some other therapies (interpersonal psychotherapy, behavioural couples therapy, counseling, brief dynamic therapy) are also recommended (with varying indications) for depression, but not for anxiety disorders.
Surveys of patients suggest that approximately twice as many patients have a preference for psychological treatment compared to medication (Kwan, Dimidjian and Rizvi, 2010). However, only a small fraction of people in the community with common mental health disorders were ever offered an evidence-based psychological treatment (McManus and Bebbington, 2009).

The English Improving Access to Psychological Therapies (IAPT) initiative was designed to address the need for much greater access to NICE recommended psychological therapies for depression and anxiety disorders (see Clark, 2011 for an overview of the programme and its history). Pilot work was undertaken in Newham and Doncaster (see Clark et al., 2009 for an evaluation) and a National Implementation Plan was published in early 2008 (Department of Health, 2008). The plan covered a period of six years during which the number of IAPT services in the country would gradually increase until all areas had a local service. Each service was required to provide NICE recommended therapy. For mild to moderate depression and several anxiety disorders (but not PTSD or social anxiety disorder) NICE recommends a stepped care model of service provision in which a substantial proportion of individuals are first offered a low intensity intervention (such as guided self-help), with individuals who fail to respond adequately to low intensity intervention being stepped up to more traditional face-to-face therapy (high intensity intervention). The IAPT services adopted this model when appropriate. Roll-out to at least 20 local services in 2008/9 was agreed for the first year. Initial progress was greater than expected with 35 services being established in that year.

Detailed outcome monitoring and ongoing evaluations of the programme are considered an integral part of IAPT. The programme stipulates a minimum dataset, which records the care provided to each service user and his or her clinical progress. High levels of pre-post data completeness are achieved by the use of a session-by-session outcome monitoring system that guarantees that a clinical endpoint is available even if a patient ends therapy earlier than expected. In July 2010, the North East Public Health Observatory published a report detailing an initial analysis of data taken from the first year of the IAPT programme (Glover, Webb and Evison, 2010). The report particularly focused on equity of access, descriptions of the treatments offered, gradings of staff and overall outcome. The report found that the overall recovery rate in the services was 42% for patients who received at least some treatment (defined as having at least 2 sessions on the assumption that the first session was always assessment). It was found that although the majority of patients received NICE compliant treatment for their disorder, a significant minority did not. However the analysis did not consider whether compliance with NICE guidance impacted on patient outcome. Significant between service variability in recovery rates was observed but predictors of this variability were also not investigated.

A recent report (Department of Health, 2012) covering the first three years of IAPT showed that roll-out of the programme remained broadly on target. In the first three years

over 150 IAPT services were established and more than one million people used the new services with an overall recovery rate in excess of 45% for those people who had completed treatment. The recovery figures are approaching the programme's 50% target, which is derived from the randomized controlled trials that generated the initial NICE recommendations (Department of Health, 2008). IAPT services' recovery rates increased year on year, with the highest recovery rates observed in the most recent time period. This increase was also seen in the number of people leaving welfare support. As with the Glover report, the IAPT three-year report did not investigate predictors of variability in outcome.

This present report takes a more detailed look at the year one IAPT data in order to identify predictors of variability in outcome at the patient level, at service level and as a function of compliance or deviation from NICE recommendations about the type of treatment that should be offered for a particular problem. The aim of this more detailed analysis is to learn lessons that can be implemented in the future to help routine clinical services enhance the outcomes that they achieve with their clientele.

Methods

Ethics statement

This study is considered a routine service evaluation. Consent was obtained from both the Department of Health and the individual services for the data analysis. Each service obtained the consent of patients for their anonymised data to be included in the Minimum Data Set for subsequent analysis.

Design

An observational, prospective cohort design. Patients who were assessed by the services were asked to complete standardized measures of depression and anxiety at every session and other measures in the Minimum Data Set (MDS: Department of Health,2011) at less frequent intervals. 32 of the 35 Year One IAPT services provided data for analysis. The remaining services were still developing their information technology systems and were unable to participate in the analysis. The data were collected between 1st October 2008 to 30th September 2009. Services varied in when they became operational. Eighteen services started collecting data in the first month, a further 10 started in the second month, the remainder started further into the year.

Patients

Up to 19,395 patients were included in the analyses. To be included they were required to satisfy a number of criteria (see Figure 1). Patients were required to have an initial assessment and to have completed their treatment by the end of September 2009 (i.e. at the end of the programme's first year). This meant that a large proportion of patients who accessed the services in year one could not be included as they had not yet completed their treatment. Patients were also required to have been clinical cases at the initial assessment and to have received as least a minimal dose of therapy. Casesness was defined as scoring above clinical/non-clinical cut-off on the depression and/or anxiety measure. To be considered someone who had at least a minimal dose of therapy, patients had to have attended at least two sessions. This was because : 1) it was thought unlikely that patients who had only one session would have received a significant amount of treatment as the first session was almost always devoted to assessment; and 2) separate pre- and post-treatment PHQ-9 and GAD-7 scores could not be collected if there was only one session. So that clinical change could be estimated patients had to have completed at least two PHQ-9 and GAD-7 questionnaires

during the course of their treatment. The percentage of treated patients that provided pairs of PHQ-9 and GAD-7 scores was unusually high for a routine cohort study. Among those who were seen at least twice and were clinical cases at initial assessment, 91.4% (20,009/21,882) had paired scores (see Figure 1). For some analyses, patients were also required to have been allocated an ICD-10 diagnosis by their service. NICE recommendations are diagnosis specific so it would not be possible to assess the impact of NICE compliance without this information. In addition, it seemed likely that overall outcome may vary with diagnosis. Finally, for some analyses patients were required to have been treated in a service that provided detailed information on the types of treatment that they received (three services were excluded for this reason, four services were excluded as they did not give patients' diagnoses and one service did not indicate whether patients were still receiving treatment or not as patients were not given an end of treatment marker). Overall, data from 24 services were included in the analysis.

Measures

Depression was assessed with the 9-item Patient Health Questionnaire Depression Scale (PHQ-9: Kroenke, Spitzer and Williams, 2001) which ranges from 0 to 27 with a recommended cut-off of 10 or above for distinguishing between clinical and non-clinical populations. Anxiety was assessed with the 7-item Patient Health Questionnaire Generalized Anxiety Disorder Scale (GAD-7: (Spitzer, Kroenke, Williams and Lo, 2006), which ranges from 0 to 21. Although the latter scale was originally developed to screen for Generalised Anxiety Disorder (GAD), it also has satisfactory (albeit lower) sensitivity and specificity for detecting other anxiety disorders when a cut off of 8 or above is used (Kroenke, Spitzer, Williams, Monahan and Löwe, 2007).

Outcome indices

Previous reports of outcomes in IAPT services have used the "recovery" index. An individual is judged to have recovered if s/he is a case at pre-treatment and has dropped below the clinical/non-clinical cut-off for depression *and* anxiety at post-treatment. This measure does not take into account whether the observed change is greater than the measurement error of the scales. As a consequence, a patient who starts treatment just above the clinical threshold and finishes treatment just below it will be classified as "recovered" even if the improvement is not statistically reliable. To get round this problem, we used a "reliable recovery" index. Patients were deemed to have reliably recovered if they scored above the clinical cut-off on the PHQ-9 and/or the GAD-7 at initial assessment, they showed reliable improvement during treatment, and they scored below the clinical cut-offs on *both* the PHQ-9 and the GAD-7 at the end of treatment. Reliable improvement was assessed using Jacobson and Truax's (1991) reliable change criteria. The measure of reliability used for the PHQ-9 and the GAD-7 was Cronbach's α, taken from the validation studies of the measures (Kroenke et al., 2001; Spitzer et al., 2006). To be considered reliable, pre-post change on the PHQ-9 needed to exceed 5.20. For the GAD-7 the comparable value was 3.53. Patients were considered to have shown reliable improvement if their PHQ-9 or GAD-7 score reliably decreased and the score for the other scale either did the same or did not reliably deteriorate. For the main predictive analyses, we focus on the reliable recovery index as this most closely corresponds to the measure normally reported by IAPT services. However, we recognize that some patients may show worthwhile improvements in therapy that fall short of full recovery. To capture this phenomenon, we report reliable improvement rates. Similarly, some patients may deteriorate during a course of therapy. To capture this phenomenon, we also report reliable deterioration rates. Patients are considered to have shown reliable deterioration if their PHQ-

9 or GAD-7 score reliably increased and the score for the other scale either did the same or did not reliably improve.

Statistical analysis

Logistic regression models were used to test whether compliance or deviation from NICE recommendations about the type of treatment that should be offered had an effect on patients' likelihood for reliable recovery. These analyses controlled for patients' initial scores on the PHQ-9 and the GAD-7 and, in the case of patients who received high intensity treatment, whether they had also received low intensity treatment prior to receiving high intensity treatment. A simultaneous entry method was used to control for the effects of the variables. NICE guidelines are diagnosis specific, so the effect of receiving, or not receiving, NICE compliant treatment was investigated within diagnostic groupings. In three diagnostic groupings sufficient patients (n>100) received treatment that was not compliant with NICE to make a comparison between compliant and non-compliant therapy possible. The groupings were: Depressive episode, Mixed Anxiety and Depressive Disorder (MADD) and Generalised Anxiety Disorder (GAD).

Logistic regression was also used to identify patient level and service level variables that predict reliable recovery. A backwards-stepwise method using the likelihood ratio was chosen as this avoids suppressive effects, and is recommended when there are no firm hypotheses (Menard, 1995). The variables initially entered in the model are shown in Table 1. A liberal criterion for selection was used (α=.2) based on the findings that conservative criteria for selection in regression analyses can lead to type II errors (Mickey and Greenland, 1989). Hosmer and Lemeshow's test was used to assess the goodness of fit of the models (Lemeshow and Hosmer, 1982).

Logistic regression analyses describe the effects of variables in terms of odds ratios. When the independent variable is dichotomous and denotes when a particular event has occurred (for example, a patient was self referred) the odds ratio is the ratio of the likelihood an event occurring in one group (self referred patients) over the odds of it occurring in the other group (non-self referred patients). When the independent variable is continuous, the odds ratio describes the increase in likelihood of a patient reliably recovering if there is a single unit increase from the mean in the independent variable i.e. if the number of sessions of treatment was found to be a significant predictor and have an odds ratio of 1.1, then for every extra session above the mean there would be a 10% increase in the likelihood of reliable recovery. These odds ratios were considered in a multivariate analysis to control for all other variables in the model.

Results

Reliable recovery, reliable improvement, and reliable deterioration

Overall, 40.3% of the 19,395 patients included in the full sample showed reliable recovery.[1] However, reliable recovery rates varied considerably from service to service, ranging from a low of 23.9% to a high of 56.5% (SD=8.0%). This can be seen in Figure 2.

Table 2 shows the reliable improvement and reliable deterioration rates. Overall, 63.7% of patients showed reliable improvement on the combination of PHQ-9 and GAD-7. As with the reliable recovery, the rates varied substantially between services, with the lowest being 43.6% and the highest being 77.1% (SD=7.1%). Psychological therapies can be

[1] A small number of patients (n=411, 2.1%) finished treatment below the clinical threshold on both the PHQ-9 and the GAD-7, but did not show reliable improvement either of these measures.

harmful as well as helpful. For this reason it is important to determine how many patients deteriorate during the course of treatment. Overall, 6.6% of patients showed reliable deterioration. Again there was considerable between service variability, with the lowest being 2.1% and the highest being 11.4% (SD=1.7%). There was a significant negative correlation between service reliable improvement rates and service reliable deterioration rates (r=-.397, p=. 027), indicating that services in which fewer people improved had a greater proportion who deteriorated.

The analyses above required patients to be cases at the start of treatment. However, a number of patients (n=3,759) started treatment below caseness, but were still seen at least twice, received some treatment, and had two scores on the PHQ-9 and the GAD-7. The proportion of these patients who showed reliable improvement was 24.3% (n=909) and the proportion of these patients who showed reliable deterioration was 11.7% (n=439). Further investigation showed that 1,024 of these patients could not show reliable improvement, as their initial scores on the PHQ-9 and GAD-7 were too low to do so (below 6 and 4, respectively). Once this has been taken into account we can see that, of the below caseness patients that could show reliable improvement, 33.2% did.

The effect of NICE compliance on reliable recovery

High intensity therapies. NICE recommends CBT as a high intensity therapy for depression and for all anxiety disorders. In the first year of the IAPT programme, the vast majority of patients were offered CBT. However, a substantial subset of patients with ICD-10 diagnoses of depressive episode, generalized anxiety disorder (GAD) or mixed anxiety and depressive disorder (MADD) received counselling. While NICE recommends counselling as well as CBT for mild to moderate depression, it does *not* recommend counselling for GAD. NICE have not released any guidance for MADD, which is technically reserved for patients with sub-threshold symptoms of anxiety and depression. However, IAPT patients diagnosed with MADD had high initial scores on the PHQ-9 (mean =16.33, SD=5.43) and the GAD-7 (mean = 14.42, SD=4.41), suggesting that many were probably best considered as individuals with both an anxiety disorder and a depressive disorder. For such individuals, current NICE guidelines would favour CBT.

To determine whether compliance with NICE guidance is associated with improved clinical outcomes, we compared the raw reliable recovery rates associated with CBT and counselling in patients with depressive episodes, GAD and MADD and also used logistic regression to control for initial symptom levels and any prior history of low intensity intervention. The reliable recovery rates for patients who received high intensity treatment and were diagnosed with a depressive episode (unadjusted for any differences in pre-treatment scores on the PHQ-9 and GAD-7) were 40.0% for those who received CBT (n=935) and 38.3% for those who received counselling (n=679). For patients diagnosed with GAD the reliable recovery rates were 54.2% for those who received CBT (n=679) and 39.7% for those who received counselling (n=302). For patients diagnosed with MADD, the respective reliable recovery rates were 39.2% (of 704 patients) and 34.4% (of 1,005 patients).

Logistic regression analyses (see Table 3) confirmed that compliance with NICE guidance was associated with higher recovery rates. Among patients who were diagnosed with a depressive episode, those who received CBT were no more or less likely to reliably recover than those who received counselling (p=.28). In contrast, among patients diagnosed with GAD, those who received CBT were 1.324 times more likely to reliably recover than those who received counselling (p< .001). Similarly, among patients who were diagnosed with MADD, patients who received CBT were 1.689 times more likely to reliably recover (p<.001). In all three logistic regression models, the data were shown to fit the model using the Hosmer and Lemeshow test (p>.05). The model for patients with a depressive episode

explained 8.4% of the variance (using Nagelkerke's R^2), and the model for patients with MADD, 6.6% and GAD, 10.7%. All models were significantly better at predicting patients' outcomes than a model than just contained a constant (p<.001).

Low intensity therapies. Self-help interventions can be offered with or without the guidance of a clinician. NICE guidelines for depression (NICE 2004b, 2009a) recommend guided self-help but not pure (non-guided) self-help. At the time of the data collection for this paper, NICE (2004a) recommended self-help for the treatment of GAD but did not mention the distinction between guided and non-guided delivery. However, in a recent revision to the GAD guideline (NICE, 2011a) guided self-help and pure self-help were both recommended, although it was acknowledged that the evidence base for pure self-help was modest.

In the year one IAPT services, the majority of patients who received low intensity interventions were given guided self-help. However, a significant sub-group had pure self-help so it was possible to assess whether compliance with NICE recommendations for low intensity treatment was also associated with enhanced reliable recovery rates. The reliable recovery rates for patients who received low intensity treatment and were diagnosed with a depressive episode (unadjusted for any differences in pre treatment scores on the PHQ-9 and GAD-7) were 38.4% for those who received guided self help (n=408) and 27.6% for those who received counselling (n=199). For patients diagnosed with MADD, the respective reliable recovery rates were 35.3% (of 388) and 35.4% (of 192) and for patients diagnosed with GAD the reliable recovery rates were 54.1% (of 279) and 52.3% (of 151).

Logistic regression analyses showed that among patients who were diagnosed with a depressive episode, those who received guided self-help were 1.561 times more likely to reliably recover than those who received pure self-help [*Wald statistic* (1) = 5.239, p= .022, *Odds ratio* = 1.561, *Lower CI* = 1.066, *Upper CI* = 2.285]. Among patients who were diagnosed with MADD and GAD, they were no more likely to reliably recover if they received pure or guided self-help. For patients with MADD the Wald statistic was 0.011 (p= .917, *Odds ratio* = 1.020, *Lower C*I = 0.698, *Upper CI* = 1.492). For patients with GAD the Wald statistic was 0.013 (p= .908, *Odds ratio* = 1.025, *Lower CI* = 0.670, *Upper CI* = 1.569). In all three logistic regression models, the data were shown to fit the model using the Hosmer and Lemeshow test (p>.05) and all models were significantly better at predicting patients' outcomes than a model than just contained a constant (p<.05). The model for patients with a depressive episode explained 10.5% of the variance (shown by Nagelkerke's R^2), the model for patients with MADD, 11.8% and the model for patients diagnosed with GAD explained 15.6 % of the variance.

For patients to be included in the analyses above they were required to have evidence that they attended IAPT services at least twice. However, some patients were provided with self-help materials in session one and were not seen again. We suspected this might be more common for people allocated to pure self-help than for people allocated to guided self-help. Further analysis confirmed that this was the case. Patients who received pure self-help were significantly less likely to have two sets of PHQ-9 and GAD-7 scores than patients who received guided self-help [X^2(1) =1024.40, p<.001, Φ=.393]. Clearly, we cannot know the outcome of these patients with any certainty. However, if we make the conservative assumption that they are unlikely to have benefited and so carry forward their session one score, the relative reliable recovery rates of patients who received guided and pure self-help are altered dramatically. This can be seen in Figure 4. Patients who received guided self – help were more likely to recover than those who received pure self-help. Logistic regression models were created to investigate whether patients who received pure self-help were more likely to reliably recover than patients who received guided self-help, if the last observation is carried forward. In all three logistic regression models, the data were shown to fit the model using the Hosmer and Lemeshow test for goodness of fit (p>.05) and all models were

significantly better at predicting patients' outcomes than a model that just contained a constant (p<.001). Receiving guided self-help remained a significant predictor of reliable recovery among patients with any of the three diagnoses investigated.

Among patients who were diagnosed with a depressive episode, those who received guided self-help (n=637) were 3.19 times more likely to recover than those who received pure self help (n=611) (*Wald statistic* (1) = 45.91, p<.001 *Odds ratio* = 3.190, *Lower CI* = 2.281, *Upper CI* = 4.462). Among patients with MADD, those who received guided self-help (n=556) were 2.595 more likely to recover than those who received pure self-help (n=596) (*Wald statistic* (1) = 32.914, p<.001 *Odds ratio* = 2.595, Lower CI = 1.873, Upper CI = 3.594). Finally, for patients diagnosed with GAD, those who received guided self help (n=358) were 2.148 times more likely to recover (n=315)(*Wald statistic* (1)= 19.015, p<.001, *Odds ratio* = 2.148, *Lower CI* = 1.523, *Upper CI* = 4.462).

The model for patients diagnosed with a depressive episode explained 13.0% of the variance (shown by Nagelkerke's R^2), the model for patients with MADD, 11.6% and the model for patients diagnosed with GAD explained 13.7% of the variance.

Of all the patients that were stepped up to high intensity intervention after a low intensity intervention, a significantly higher proportion had received pure self-help than guided self-help [$X^2(1)$=466.09, p<.001, Φ=.287]. The proportion of patients who were stepped up after receiving guided self-help was 25.7%, compared to 54.5% of patients who received pure self-help. This finding would appear to confirm the inferiority of pure self-help.

Factors predicting reliable recovery

A logistic regression was used to investigate the patient and service level factors that predict reliable recovery. As mentioned earlier, this logistic regression focused on the subset of patients (n=11, 535) who had been given an ICD-10 diagnosis by their service and for whom the relevant service level variables were available[2]. The reliable recovery rate in this sample (40.3%) is essentially the same in the full sample, as were the proportions of patients who showed reliable improvement (64.6%) and reliable deterioration (6.8%). The model was shown to fit the data well, as Hosmer and Lemeshow's test was non-significant [X^2(8) =4.698, p=.789]. Nagelkerke's R^2 showed that the model explained 13.2% of the variance. The model differed significantly from a model which only included the constant [X^2(14) =1188.521, p<.001]. The model successfully identified 81.4% of patients who did not reliably recover and 41.9% of those who did. Overall, the model correctly identified 65.5% of patients' outcomes. Table 1 shows the patient and service level variables that were investigated and Table 4 shows those variables that were significant predictors of reliable recovery.

Patient level variables

Initial Severity: Patients' initial PHQ-9 and GAD-7 scores had a significant effect on reliable recovery. Patients with higher initial scores were less likely to reach reliably recovery. However, this does not mean that patients with moderate or severe symptoms benefited less from therapy that those with mild symptoms. Indeed there was some evidence to the contrary (see Figure 3). Patients were divided into three initial severity groupings (moderate, moderately severe or severe) on the basis of published norms (Kroenke et al., 2001). A Kruskal Wallis test comparing change scores found that increasing severity was associated with greater improvement [X^2(2) =457.64, p<.001]. The mean change for patients initially classed as having moderate depressive symptomatology on the PHQ-9 was 4.47 (SD=5.35) in

[2]Most of the variables that are significant predictors in this logistic regression were also significant in a logistic regression that was run on the full sample for sensitivity purposes. However, the model fit was less good, perhaps because ICD-10 diagnosis, which is a significant predictor, could not be included.

comparison to 7.99 (SD=7.63) for patients with severe depressive symptomatology. The same pattern of results was found for patients' scores on the GAD-7 when patients' initial scores were classed as mild, moderate or severe based on published norms (Spitzer et al., 2006). The mean change on the GAD-7 for patients initially classed as 'mild' on the measure was 2.16 (SD=4.32) in comparison to 6.77 (SD=6.27) for patients classed as 'severe'. All groups showed marked improvement with the greatest improvement being shown by the patients who started treatment with higher scores [$X^2(2) = 1244.01$, $p<.001$].

Self-referral: Whether or not patients referred themselves to treatment was not a predictor of reliable recovery. However, patients who reliably recovered and had self-referred had fewer therapy sessions than patients who reliably recovered and were not self-referred [Mann-Whitney U=1932729, p=.005, r=.031]. This suggests that the process of self-referral may facilitate therapy.

Treatment received: The model shows that when all things are considered, reliable recovery is was less likely if patients received 'other treatment' and was more likely if patients received high intensity treatment, compared to not receiving these treatments. "Other" treatment was a code used when the intervention was not a recognized high or low intensity intervention.

Diagnosis: Patients who were diagnosed with depressive episode, GAD, MADD or PTSD were significantly more likely to recover than patients who did not receive these diagnoses.

Service level variables

Use of Stepped Care: Patients treated in services in which a greater proportion of patients who received low intensity treatment were stepped up from low intensity to high intensity care had higher overall rates of reliable recovery.

Mean number of therapy sessions: Patients treated in services with higher average numbers of therapy sessions were more likely to recover than patients treated in services in which fewer sessions were offered. This finding was significant for low intensity therapy and for patients who received any high intensity treatment.

Staff Salary Bands: In the National Health Service (NHS) staff receive remuneration based on a national standardised pay-scale. This is negotiated centrally and is organised into 'Agenda for Change' (AfC) bands, which range from Band 1 to Band 9. The distribution of staff salary bandings within a service was a significant predictor of reliable recovery. Patients treated in services where a greater proportion of therapist sessions were undertaken by therapists banded at AfC band 7 or above, were more likely to reliably recover than patients treated in services where a smaller proportion of sessions were undertaken by such workers. In year one of the IAPT programme most trainee therapists would have been paid below AfC Band 7. This finding may therefore indicate that services with a larger cohort of clinically active experienced staff achieve higher reliable recovery rates.

Size of the service: The number of patients treated at a service was found to be an important predicting factor in patients' reliable recovery. The greater the number of patients treated at the service, the more likely it was that patients treated at the service would reliably recover.

Discussion

The English Improving Access to Psychological Therapies (IAPT) initiative is probably the world's largest single programme for disseminating evidence-based psychological therapies to a general population. The use of a session-by-session outcome monitoring system has

ensured that clinical outcomes are recorded on almost everyone who is treated in the IAPT services. This has brought unparalleled public transparency to mental health provision with key performance indicators (KPIs) for all IAPT services published every three months on the national Health and Social Care Information Centre website (http://www.hscic.gov.uk/mentalhealth). The large database also provides an opportunity to learn lessons about the way in which psychological therapies might best be provided in order to maximize clinical outcomes. This paper, which focuses on data from the first year of the programme, is one of the first attempts to identify such lessons. In future years further analyses based on the evolving database will be published. The present analyses were conducted in close collaboration with the IAPT clinical services. Regional representatives suggested questions that could be investigated in the analysis. Several key findings have emerged from the analyses that are likely help the local services further develop their provision in the future. Many may also be helpful for commissioners and clinicians in other countries as they pursue their own plans to increase the availability of evidence-based psychological therapies for their own populations.

Key performance indicators

Support for the IAPT programme critically depends on it being able to show that it can achieve the kind of results one might expect from published randomized controlled trials of psychological therapies. For this reason, IAPT services were all asked to report a simple measure of outcome from the beginning of the initiative. The measure was "recovery", which was judged to have occurred if a patient scored above the clinical cut-off on the PHQ-9 and/or the GAD-7 at pre-treatment and scored below the clinical cut-off on *both* at discharge from the service. This measure, which we will term the recovery index is easy to calculate and has served the programme well. However, it has several limitations.

First, the recovery index does not take into account the measurement error associated with each scale so it is possible that some mild cases will be classified as recovered when the observed symptom reduction is not reliable. To get round this problem, we used a modified, reliable recovery index (RRI) in our analyses and recommend that it is used in the future. Encouragingly, the overall findings with the reliable recovery index (40.3% of patients classified as recovered) are not much different from those for the original recovery index (42.4 % of patients classified as recovered). However, it is possible that in some services the difference will be larger and it would be important to know this.

Second, the binary nature of the recovery index means that no information is provided on the improvements that patients who did not fully recover may have made during treatment. It was suspected that many patients who had not fully recovered might still have made worthwhile gains. The adoption of a reliable improvement measure has demonstrated that is in fact what happened. While 40.3% of patients who were initial cases showed reliable recovery, 63.7% showed reliable improvement.

Third, the recovery index provides no information about deterioration. Psychological therapies have the potential to do harm as well as good. Given this point, it is important to assess the extent to which patients may get worse during a course of therapy. The reliable deterioration measure reported here indicated that 6.6% of patients got worse during their treatment in IAPT services. This overall rate is probably less than one would observe in a population allocated to a wait-list and so is probably not a cause for concern. However, it may be higher in some services and it should be carefully monitored in the future. Services may also wish to conduct their own audits of individuals who show reliable deterioration in order to identify any patterns (particular subsets of individuals, therapists, or treatments) that can inform further service development.

Finally, the main logistic regression found that RRI rates were highest in patients who had an ICD-10 diagnosis of depressive episode, GAD, MADD or PTSD. It is unclear how one should interpret this finding. It may mean that people with these conditions show greatest benefit from IAPT treatment. However, it could also be an artifact of unknown variation in natural recovery rates or a quirk of the measurement system used in year one. In connection with the latter, the PHQ-9 and the GAD-7 are sensitive measures for detecting and assessing change in depression and GAD but are less sensitive for other anxiety disorders, such as social anxiety disorder and obsessive-compulsive disorder. A revised IAPT minimum dataset has now been published (Department of Health, 2011) which includes sensitive measures of these conditions so future analyses will be able to investigate this issue.

Importance of compliance with NICE's recommendations for treatment types

A defining feature of the IAPT programme is that it aims to greatly increase the availability of NICE recommended psychological therapies for anxiety disorders and depression (Department of Health, 2008). In line with this aim, most of the patients treated in the first year of the programme received a NICE recommended treatment. However, for three disorders (depression, generalized anxiety disorder, mixed anxiety and depressive disorder) a significant minority received an intervention that is not recommended by NICE. This created a natural experiment in which it was possible to assess whether deviation from NICE guidelines was associated with reduced reliable recovery rates. The main analysis of the importance of compliance with NICE guidance focused on individuals who had been ascribed an ICD-10 diagnosis by their service, as NICE guidelines are diagnosis specific.

When considering high intensity treatments, NICE (2005a, 2005b, 2009a, 2011a, 2013) recommends both CBT and counseling for mild to moderate depression but only recommends CBT for any of the anxiety disorders. The observed results were in line with these recommendations. In particular, CBT and counseling were associated with similar reliable recovery rates in depression but CBT was associated with significantly higher reliable recovery rates than counseling in generalized anxiety disorder (GAD) and in mixed anxiety and depressive disorder (MADD). In depression, there was no difference in recovery rates between CBT and counseling. However in GAD and MADD patients who received CBT were more likely to recover than those who received counseling.

Turning to low intensity treatment, for depression NICE (2004b,2009a) recommends guided self-help but not pure self-help. The observed results were in line with this recommendation. Reliable recovery rates were significantly higher among those who received guided self-help than among those who received pure self-help. The same pattern was also observed among patients with MADD. For GAD, NICE guidelines are less clear. The original guideline (NICE 2004b) failed to distinguish between guided and pure self-help and the revised guideline (2011a) recommends both, while acknowledging that the evidence base for pure self-help is modest. Our findings are similarly unclear. If one looks at those individuals with GAD who were seen at least twice in the services, there is no difference in reliable recovery rates between guided and pure self-help. However, a significantly greater proportion of people who were given pure self-help were only seen once. We cannot know how these people faired but if one assumes no benefit, then the overall reliable recovery rate is significantly lower for pure self-help than guided self-help. This result raises concern about the use of pure self-help in GAD and, at the least, suggests that if services choose to use pure self-help, they should give patients a follow-up appointment to assess whether any benefit has occurred and to move patients onto an assisted, low or high intensity treatment if there is no improvement.

A further indication of the importance of compliance with NICE guidance concerns the findings with respect to "other" treatment in the logistic regression that included patients

who had not been given an ICD-10 diagnosis as well as those for whom a diagnosis was available. The "other" category was reserved for treatments that were not in the list of treatments that would be recommended by NICE for any of the disorders covered by the IAPT programme. In line with NICE's recommendations, "other" treatment was associated with a lower overall reliable recovery rate.

The comparisons above between NICE compliant and non-compliant treatments are naturalistic. The logistic regressions controlled for initial severity. However, as patients were not randomized to the different types of treatment it is always possible that there some unobserved, but systematic differences between individuals who received the NICE complaint and non-compliant treatments were present. Given this point, it would be wrong to take our findings as a demonstration of efficacy *per se*. Instead they simply indicate that when one looks at treatments naturalistically deployed in the field, the pattern of results that is obtained is largely in line with what one might expect given NICE guidance.

As the IAPT programme has developed it has expanded patient choice among NICE recommended treatments for depression. In addition to counseling, couples therapy, interpersonal psychotherapy and brief psychodynamic therapy are all now available in some IAPT services and it is estimated by the Department of Health (2013) that around 30% of IAPT high intensity therapists are able to deliver these non-CBT treatments. Future analyses of IAPT databases will no doubt investigate whether these different treatments are associated with similar outcomes among depressed patients.

Patient and service level predictors of reliable recovery

Several patient and service level variables were found to be significant predictors of reliable recovery.

Initial severity: Patients whose initial symptom severity is moderate to severe need to show considerably more symptomatic improvement than patients with mild to moderate symptoms in order to be classified as reliably recovered. Given this point, it is perhaps not surprising that initial severity predicted reliable recovery. However, analysis of continuous change scores indicated that patients with more severe symptoms showed as much, indeed slightly more, symptomatic improvement than those with mild symptoms. The greater change in more severe patients may be regression to the mean. However, the fact that substantial change was shown at all levels of initial severity suggests that IAPT services benefit patients over the full range of severity. It also raises the question of whether in the future the key performance indicators should be expanded to include an index that more accurately captures the amount of improvement that a patient makes independent of start level. Pre-treatment to post-treatment effect size would seem an obvious candidate.

Self-referral: Traditionally the English NHS has restricted access to specialist services to individuals who are referred by their general practitioner (GP). However, the IAPT programme allows self-referral because there was concern that some patients with depression and/or anxiety disorders may be reluctant to contact their GP in the first place (Department of Health, 2008) and one of the pilot sites (Newham) found that individuals from the black and ethnic minority community and some anxiety disorders were under-represented in GP referrals (Clark et al., 2009). As in the analysis of the original pilot sites (Clark et al., 2009), patients who were treated in the first year of the national roll-out did not differ in their recovery rates as a function of how they were referred. However, it is interesting to note that self-referred patients who showed reliable recovery had received less treatment sessions than GP referred patients who also achieved reliable recovery. Anecdotally, it seems that self-referred patients are more likely to have sought out detailed information about the services (from websites, leaflets etc.) in advance of their first

appointment. This may help ensure that they are more engaged in treatment from the start. Further research could helpfully explore this possibility.

Use of stepped care: Stepped care is at the heart of the IAPT clinical model. With the exception of patients with PTSD or social anxiety disorder, it is suggested that patients with mild to moderate symptoms of depression or other anxiety disorders could be offered low intensity (such as guided self-help) interventions initially, with patients who fail to recover at that level being stepped up to high intensity intervention (Department of Health, 2010, p32). The finding that services that have an overall higher step-up rate also have an overall higher reliable recovery rate suggests that it is important that services make full use of their stepped care system and encourage patients to continue from low to high intensity work, if appropriate.

Mean number of therapy sessions: In addition to specifying certain types of therapy, NICE also provides recommendations about the number of therapy sessions that patients should be offered. In general, it is recommended that patients should be offered *up to* the number of sessions provided in the randomized controlled trials that generated the relevant NICE guideline. For high intensity treatments this would generally be in the range of 12-20 sessions, depending on diagnosis and severity. Our finding that services that offered higher mean numbers of low intensity and high intensity treatment sessions had overall higher recovery rates would seem to support NICE's position.

Staff salary bands: IAPT aims to increase access to evidence-based psychological therapies by expanding the work-force that is trained to deliver such treatments. In the first year, most of the staff in the service were still in training. The Department of Health (2008) recommended that IAPT services should have a core of at least a third of their staff who were already fully trained in order to provide supervision to trainees and treat the more complex cases themselves. The finding that overall rates of reliable recovery were higher in services in which a larger number of therapy sessions were provided by staff in salary bands (AfC7 or above) that are usually reserved for experienced staff would appear to support this recommendation. If this is the correct interpretation of the finding, the relationship between a service's distribution of salary bands and its overall reliable recovery rate may change in future years as services increase the number of fully trained low intensity workers that they employ. This is because fully trained low intensity workers would normally be employed at AfC 5 or 6.

Size of the service: The average number of patients treated per day in a service was related to the overall outcome. Services that treated larger numbers of patients had higher overall reliable recovery rates. At this stage it is unclear how to interpret this finding. One possibility is that, on average, higher volume services have more clearly developed organizational procedures (including supervision protocols) and had more practice with the IAPT model. However, there were no direct measures of these concepts and further research is required to clarify the benefits of larger services.

Limitations

This study has a number of limitations, most obviously that it was not a controlled experiment, and therefore the results should not be treated as such. However, the study does allow us to see whether the results from randomized controlled trials can be implemented in routine care on a national scale and to identify what factors in that routine care might affect outcome. A limitation to the analysis of variation between services was that the service variables were derived from patient level variables. This method has an advantage as it creates a composite picture of the service over the course of a year. However, it is also a disadvantage as the analyses treat operationally dynamic variables as static across the period of a year. Services may have changed their policies over the course of the year, as services'

policies change over time a better understanding of the impact in varying services' policies may be gained. Although IAPT services were good at collecting session by session outcome data, they were less good at giving patients' provisional diagnoses. This limited the sample used in the predictor analyses. However, sensitivity analyses conducted to investigate whether the patient and service level factors that predicted improved reliable recovery rates generalised to the full sample that included patients who had not received a diagnosis suggest that the sample restriction was not a serious problem.

Implications for practice

The study has two broad implications for the design and management of routine psychological therapy services. First, the use of a session by session outcome monitoring system made it possible to obtain high levels of pre-treatment to post-treatment data completeness (over 91% of cases). High levels of data completeness are important as in a previous study of a routine service (Clark et al. 2009) we found that patients who failed to provide post-treatment outcome data tended to have done less well. Adopting a session-by-session outcome monitoring system might enable services with low data completeness rates to improve their completeness rates and so obtain a more accurate picture of the benefits of the service that they provide. Second, the patient and service level characteristics that predicted higher reliable recovery rates in our study are generally consistent with NICE recommendations and the IAPT model, some aspects of which may be of interest to policy makers, commissioners and clinicians in other countries as they strive to improve access to psychological therapies within their own healthcare systems. Further information on the IAPT programme, including its recommended quality standards for psychological therapy services can be found on the IAPT website (www.iapt.nhs.uk).

Table 1. Possible predictors of reliable recovery included in the logistic regression

Patient Level Variables	Service Level Variables
Initial PHQ-9 scores	Service Salary Banding Distribution[3]
Initial GAD-7 scores	Service Self-Referral
Whether or not patients were self-referred	The median number of sessions given by the service to patients who received low intensity treatment only
Whether the patient received the low intensity therapy	The median number of sessions given by the service to patients who received high intensity treatment either on its own or after receiving low intensity treatment
Whether the patient received the high intensity therapy	Service Size (The number of patients treated at the service divided by the length of time a service was operating for.)
Whether the patient received both low and high intensity therapy	Proportion of patients who received low intensity treatment who also received high intensity
Whether the patient received 'other treatment'	

[3] The effect, if any, of therapist salary banding on patient recovery was investigated using the logistic regression model. In order to do this, some preliminary analysis was required to determine the most appropriate cutting point. We calculated the relationship between the overall reliable recovery rates for services and the proportion of therapy sessions that were delivered by therapists at Agenda for Change band X and above, where X ranged from 4 to 8d (the highest banding a clinician in IAPT services could have). The strongest relationship in this sample was observed when X was 7, so this was chosen as the Agenda for Change cutting point for the logistic regression analysis.

Table 2. The proportions of the patients who showed reliable deterioration, no reliable change or reliable improvement on the PHQ-9 and the GAD-7

		Reliable Change Measured on GAD-7		
		Reliable Deterioration	No Reliable Change	Reliable Improvement
Reliable Change Measured on the PHQ-9	Reliable Deterioration	*1.2% (n=241)*	*1.7% (n=337)*	*0.2% (n=44)*
	No Reliable Change	*3.7% (n=711)*	*29.0% (n=5,617)*	*16.8% (n=3,262)*
	Reliable Improvement	*0.4% (n=84)*	*7.5%(n=1,445)*	*39.5% (n=7,654)*

Table 3. Summary of logistic regression models investigating whether receiving CBT or counselling has an impact on patients' likelihood of reliable recovery

Diagnosis Group	Variable	B	S.E.	Wald	Sig.	Odds Ratio	95% C.I. for Odds ratio	
							Lower	Upper
Depressive Episode	Initial PHQ-9 Scores	-0.073	0.012	36.38	<.001	0.929	0.907	0.952
	Initial GAD-7 Scores	-0.048	0.013	13.488	<.001	0.953	0.929	0.978
	Patient was Stepped Up	-0.183	0.106	2.99	0.084	0.833	0.677	1.025
	Patient received CBT (in comparison to patients who received counselling)	0.116	0.107	1.177	0.278	1.124	0.91	1.387
	Constant	1.271	0.205	38.639	<.001	3.566		
MADD	Initial PHQ-9 Scores	-0.055	0.011	24.316	<.001	0.947	0.927	0.968
	Initial GAD-7 Scores	-0.046	0.014	10.578	0.001	0.956	0.93	0.982
	Patient was Stepped Up	0.186	0.104	3.164	0.075	1.204	0.981	1.478
	Patient received CBT (in comparison to patients who received counselling)	0.281	0.106	6.973	0.008	1.324	1.075	1.632
	Constant	0.942	0.198	22.598	<.001	2.564		
GAD	Initial PHQ-9 Scores	-0.067	0.013	26.485	<.001	0.935	0.912	0.959
	Initial GAD-7 Scores	-0.055	0.018	9.831	0.002	0.947	0.915	0.98
	Patient was Stepped Up	0.186	0.104	3.164	0.075	1.204	0.981	1.478
	Patient received CBT (in comparison to patients who received counselling)	0.524	0.149	12.377	<.001	1.689	1.261	2.263
	Constant	1.303	0.262	24.758	<.001	3.68		

Table 4. Variables included in the model after stepwise removal

Variable	B	S.E.	Wald	Sig.	Odds ratio	95% C.I. for Odds ratio	
						Lower	Upper
Patient Level Variables							
Initial PHQ-9 Scores	-0.08	0.004	342.904	<.001	0.923	0.915	0.931
Initial GAD-7 Scores	-0.047	0.005	82.125	<.001	0.954	0.945	0.964
Patient received high intensity treatment	0.109	0.046	5.553	0.018	1.116	1.019	1.222
Patient received 'other treatment'	-0.393	0.135	8.445	0.004	0.675	0.518	0.88
Depressive Episode Diagnosis	0.184	0.068	7.277	0.007	1.202	1.052	1.373
MADD Diagnosis	0.146	0.068	4.599	0.032	1.157	1.013	1.322
GAD Diagnosis	0.369	0.074	25.023	<.001	1.447	1.252	1.672
Phobias Diagnosis	0.167	0.109	2.352	0.125	1.182	0.955	1.463
PTSD Diagnosis	0.381	0.158	5.837	0.016	1.464	1.075	1.995
Service Level Variables							
Step Up Rate	1.074	0.128	70.603	<.001	2.926	2.278	3.758
Median number of sessions given to patients who received low intensity treatment	0.186	0.027	47.938	<.001	1.204	1.142	1.269
Median number of sessions given to patients who received high intensity *(either alone or after being stepped up)*	0.069	0.02	12.43	<.001	1.071	1.031	1.113
Proportion of sessions undertaken by therapists banded at AfC 7 or above	0.631	0.211	8.945	0.003	1.880	1.243	2.844
Size of service *(Number of Patients completing treatment at the service per day)*	0.162	0.024	47.105	<.001	1.176	1.123	1.231
Constant	-0.499	0.179	7.744	0.005	0.607		

Figure 1. Cohort used in the analyses

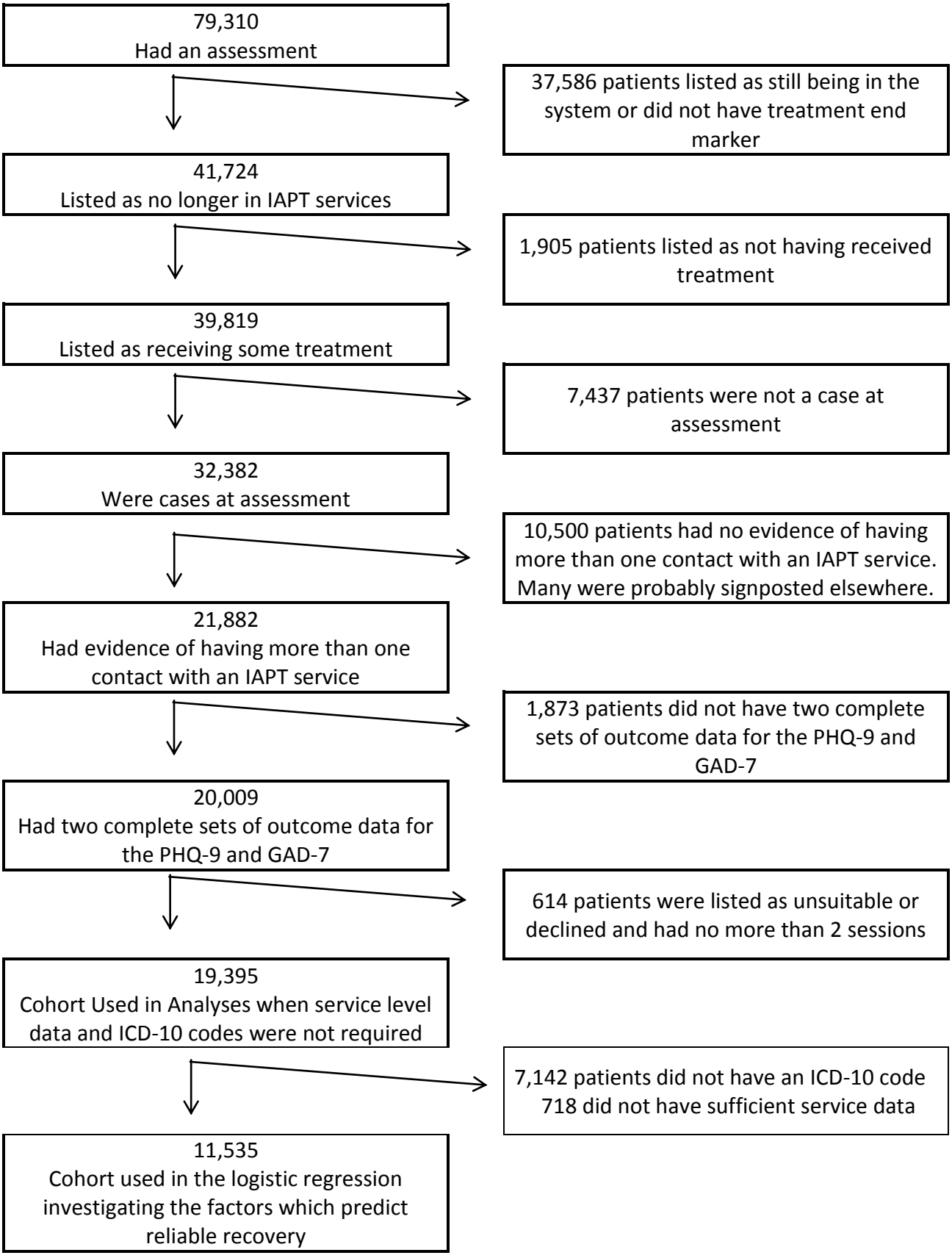


Figure 2. Variability in service reliable recovery rates

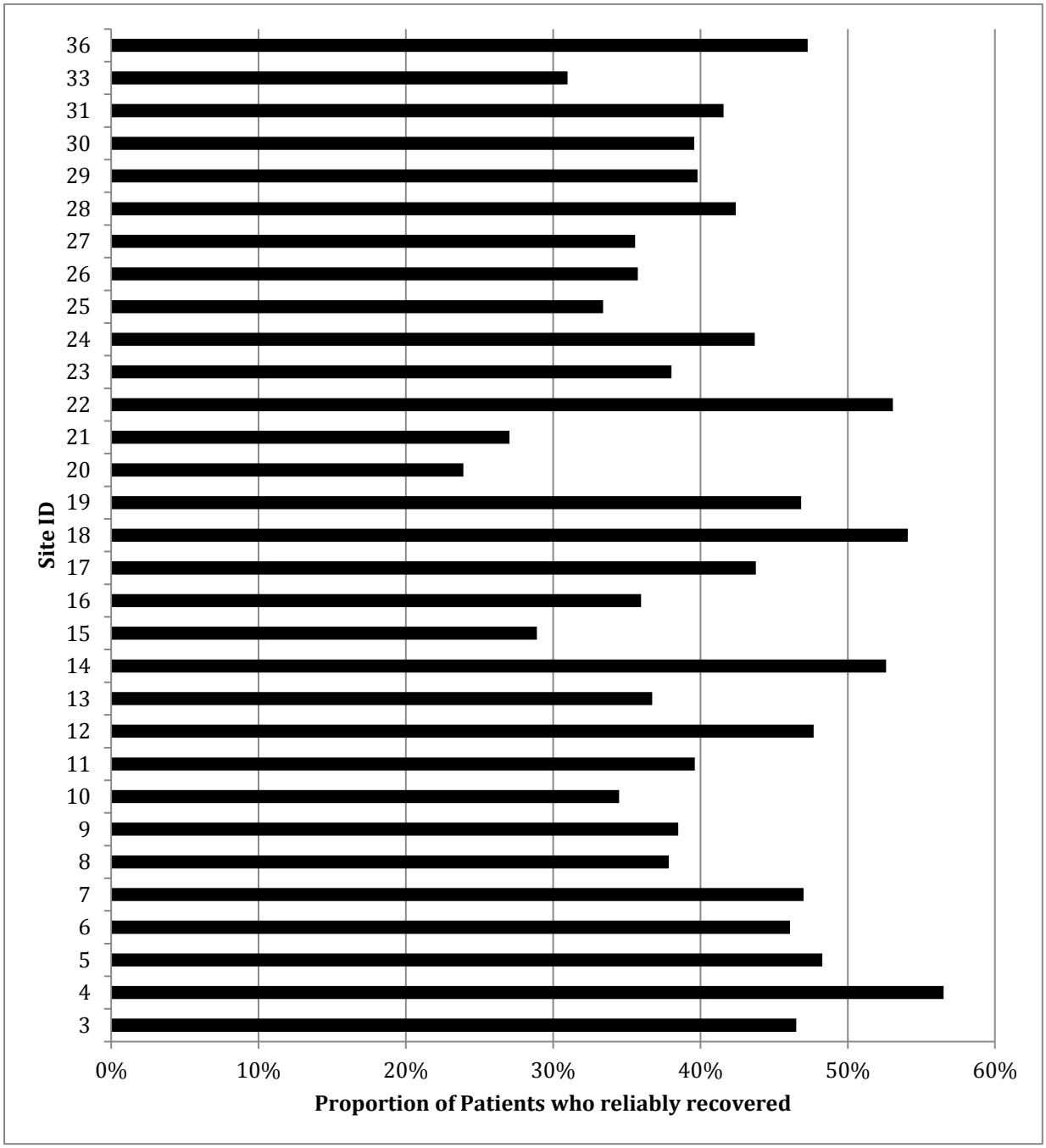


Figure 3. Change in PHQ-9 scores as a function of initial severity

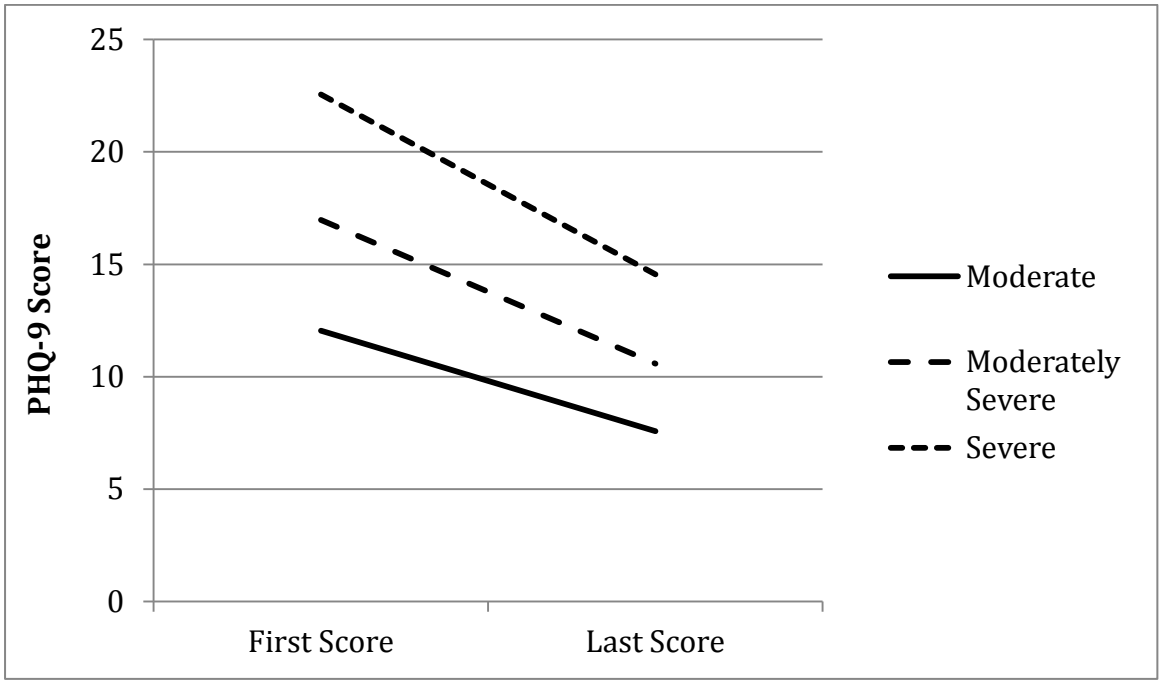
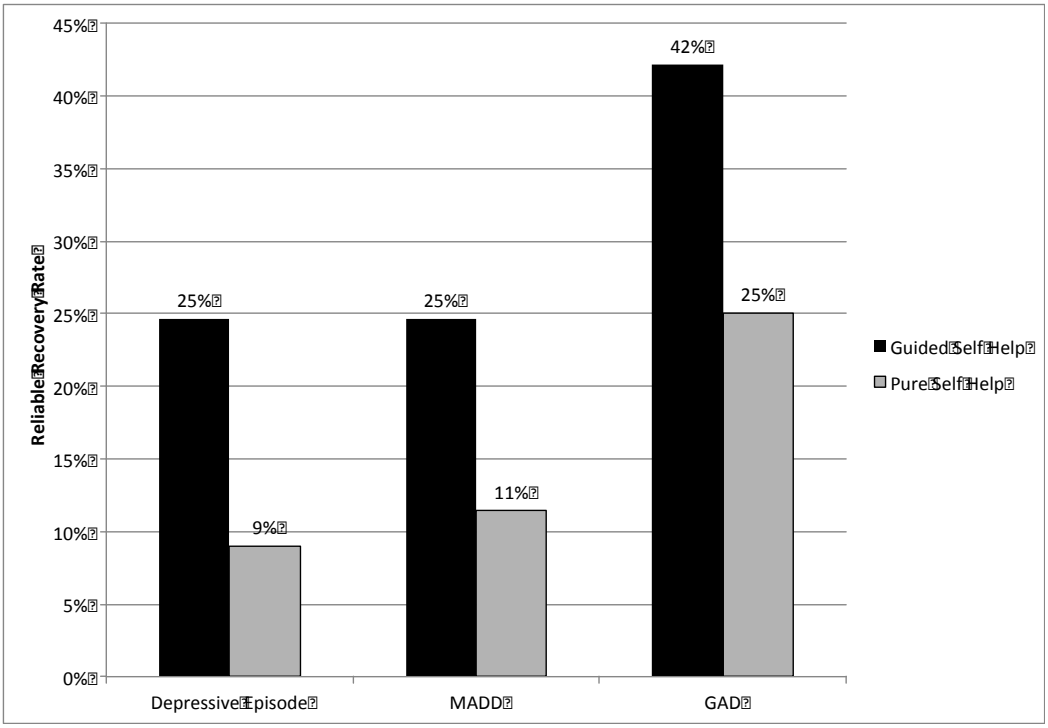


Figure 4. Comparison of Reliable Recovery Rates between patients who received guided and pure self-help by diagnosis among patients whose last scores on the PHQ-9 and GAD-7 scores are carried forward



References:

Clark, D. M. (2011). Implementing NICE Guidelines for the Psychological Treatment of Depression and Anxiety Disorders: The IAPT Experience. *International Review of Psychiatry, 23*, 375-384.

Clark, D. M., Layard, R., Smithies, R., Richards, D. A., Suckling, R. and Wright, B. (2009). Improving Access to psychological Therapy: Initial Evaluation of Two UK Demonstration Sites. *Behaviour Research and Therapy, 47*(11), 910-920. doi: 10.1016/j.brat.2009.07.010

Department of Health (2008). *IAPT Implementation Plan: National Guidelines for Regional Delivery*. Author: UK. Available at www.iapt.nhs.uk.

Department of Health (2010). *Realising the Benefits: IAPT at Full Roll-Out*. Author: UK. Available at www.iapt.nhs.uk.

Department of Health (2011). *The IAPT Data Handbook. Version 2.0.1*. Author: UK. Available at www.iapt.nhs.uk

Department of Health (2013). *IAPT Three Year Report: The First Million Patients*. Author: UK. Available at www.iapt.nhs.uk

Ekers, D., Richards, D. and Gilbody, S. (2008). A Meta-Analysis of Randomized Trials of Behavioural Treatment of Depression. *Psychological Medicine, 38*(5), 611-624.

Glover, G., Webb, M. and Evison, F. (2010). *Improving Access to Psychological Therapies: A Review of Progress Made by Sites in the First Roll-Out Year*. North East Public Health Observatory. (Available from www.iapt.nhs.uk)

Jacobson, N. S. and Truax, P. (1991). Clinical Significance : A Statistical Approach to Denning Meaningful Change in Psychotherapy Research. *Psychology, 59*(1), 12-19.

Kroenke, K., Spitzer, R. L., Williams, J., Monahan, P. O. and Löwe, B. (2007). Anxiety Disorders in Primary Care: Prevalence, Impairment, Comorbidity, and Detection. *Annals of Internal Medicine, 146*(5), 317.

Kroenke, K., Spitzer, R. L. and Williams, J. B. W. (2001). The PHQ-9. *Journal of General Internal Medicine, 16*(9), 606-613.

Kwan, B. M., Dimidjian, S. and Rizvi, S. L. (2010). Treatment Preference, Engagement, and Clinical Improvement in Pharmacotherapy Versus Psychotherapy for Depression. *Behaviour research and therapy, 48*(8), 799-804.

McManus, S., Bebbington, P., National Centre for Social, R., University of Leicester. Dept. of Health, S. and Great Britain. National Health Service. Information, C. (2009). *Adult Psychiatric Morbidity in England, 2007 : Results of a Household Survey*. (9781904599944 190459994X). London: NHS Information Centre.

NICE (2004a). *Anxiety: Management of Anxiety (Panic Disorder, With and Without Agoraphobia, and Generalised Anxiety Disorder) in Adults in Primary, Secondary and Community Care*. Clinical Guideline 22. London: National Institute for Health and Clinical Excellence. Available at www.nice.org.uk.

NICE (2004b). *Depression:Management of Depression in Primary and Secondary Care*. Clinical Guideline 23. London: National Institute for Health and Clinical Excellence. Available at www.nice.org.uk.

NICE (2005a). *Obsessive – Compulsive Disorder:Core Interventions in the Treatment of Obsessive – Compulsive Disorder and Body Dysmorphic Disorder*. Clinical Guideline 31. London: National Institute for Health and Clinical Excellence. Available at www.nice.org.uk.

NICE (2005b). *Post-Traumatic Stress Disorder (PTSD): The Management of PTSD in Adults and Children in Primary and Secondary Care*. Clinical Guideline 26. London: National Institute for Health and Clinical Excellence. Available at www.nice.org.uk.

NICE (2006). *Computerized Cognitive Behaviour Therapy for Depression and Anxiety*. Technology Appraisal 97 . London: National Institute for Health and Clinical Excellence. Available at www.nice.org.uk.

NICE (2009a). *Depression: Treatment and Management of Depression in Adults*. Clinical Guideline 90. London: National Institute for Health and Clinical Excellence. Available at www.nice.org.uk.

NICE (2009b). *Depression in Adults with a Chronic Physical HealthProblem: Treatment and Management*. Clinical Guideline 91. London: National Institute for Health and Clinical Excellence. Available at www.nice.org.uk.

NICE (2011a). *Generalised Anxiety Disorder and Panic Disorder (With or Without Agoraphobia) in Adults: Management in Primary, Secondary and Community Care*. Clinical Guideline 113. London, UK: National Institute for Health and Clinical Excellence. Available at www.nice.org.uk.

NICE (2011b). *Common Mental Health Disorders: Identification and Pathways to Care*. Clinical Guideline 123. London, UK: National Institute for Health and Clinical Excellence. Available at www.nice.org.uk.

NICE (2013). *Social Anxiety Disorder: Recognition, Assessment and Treatment*. London, UK: National Institute for Health and Clinical Excellence. Available at www.nice.org.uk

Spitzer, R. L., Kroenke, K., Williams, J. B. W. and Lo, B. (2006). A Brief Measure for Assessing Generalized Anxiety Disorder. *Response, 166*, 1092-1097.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1226	Stephen Gibbons Sandra McNally	The Effects of Resources Across School Phases: A Summary of Recent Evidence
1225	Cornelius A. Rietveld David Cesarini Daniel J. Benjamin Philipp D. Koellinger Jan-Emmanuel De Neve Henning Tiemeier Magnus Johannesson Patrik K.E. Magnusson Nancy L. Pedersen Robert F. Krueger Meike Bartels	Molecular Genetics and Subjective Well-Being
1224	Peter Arcidiacono Esteban Aucejo Patrick Coate V. Joseph Hotz	Affirmative Action and University Fit: Evidence from Proposition 209
1223	Peter Arcidiacono Esteban Aucejo V. Joseph Hotz	University Differences in the Graduation of Minorities in STEM Fields: Evidence from California
1222	Paul Dolan Robert Metcalfe	Neighbors, Knowledge, and Nuggets: Two Natural Field Experiments on the Role of Incentives on Energy Conservation
1221	Andy Feng Georg Graetz	A Question of Degree: The Effects of Degree Class on Labor Market Outcomes
1220	Esteban Aucejo	Explaining Cross-Racial Differences in the Educational Gender Gap
1219	Peter Arcidiacono Esteban Aucejo Andrew Hussey Kenneth Spenner	Racial Segregation Patterns in Selective Universities
1218	Silvana Tenreyro Gregory Thwaites	Pushing On a String: US Monetary Policy is Less Powerful in Recessions
1217	Gianluca Benigno Luca Fornaro	The Financial Resource Curse
1216	Daron Acemoglu Ufuk Akcigit Nicholas Bloom William R. Kerr	Innovation, Reallocation and Growth

1215	Michael J. Boehm	Has Job Polarization Squeezed the Middle Class? Evidence from the Allocation of Talents
1214	Nattavudh Powdthavee Warn N. Lekfuangfu Mark Wooden	The Marginal Income Effect of Education on Happiness: Estimating the Direct and Indirect Effects of Compulsory Schooling on Well-Being in Australia
1213	Richard Layard	Mental Health: The New Frontier for Labour Economics
1212	Francesco Caselli Massimo Morelli Dominic Rohner	The Geography of Inter-State Resource Wars
1211	Stephen Hansen Michael McMahon	Estimating Bayesian Decision Problems with Heterogeneous Priors
1210	Christopher A. Pissarides	Unemployment in the Great Recession
1209	Kevin D. Sheedy	Debt and Incomplete Financial Markets: A Case for Nominal GDP Targeting
1208	Jordi Blanes i Vidal Marc Möller	Decision-Making and Implementation in Teams
1207	Michael J. Boehm	Concentration versus Re-Matching? Evidence About the Locational Effects of Commuting Costs
1206	Antonella Nocco Gianmarco I. P. Ottaviano Matteo Salto	Monopolistic Competition and Optimum Product Selection: Why and How Heterogeneity Matters
1205	Alberto Galasso Mark Schankerman	Patents and Cumulative Innovation: Causal Evidence from the Courts
1204	L Rachel Ngai Barbara Petrongolo	Gender Gaps and the Rise of the Service Economy
1203	Luis Garicano Luis Rayo	Relational Knowledge Transfers
1202	Abel Brodeur	Smoking, Income and Subjective Well-Being: Evidence from Smoking Bans

The Centre for Economic Performance Publications Unit
Tel 020 7955 7673 Fax 020 7404 0612
Email info@cep.lse.ac.uk Web site http://cep.lse.ac.uk