An evaluation of a new service model: Improving Access to Psychological Therapies demonstration sites 2006-2009

Glenys Parry[1], Michael Barkham[1], John Brazier[1], Kim Dent-Brown[1], Gillian Hardy[1], Tony Kendrick[2], Jo Rick[3], Eleni Chambers[1], Tom Chan[4], Janice Connell[1], Rebecca Hutten[1], Simon de Lusignan[5], Clara Mukuria[1], Dave Saxon[1], Peter Bower[3] and Karina Lovell[3]

[1] University of Sheffield
[2] Hull York Medical School, University of Hull
[3] University of Manchester
[4] St George's Medical School, University of London
[5] University of Surrey

Published June 2011

Address for correspondence:

Professor Glenys Parry
Professor of Applied Psychological Therapies
Centre for Psychological Services Research
ScHARR, Regent Court
30 Regent Street
SHEFFIELD
S1 4DA
United Kingdom

Email: G.D.Parry@sheffield.ac.uk

Contents

List of tables

List of figures

Glossary of terms/abbreviations

A&E	Accident and Emergency department
ANCOVA	Analysis of co-variance
ANOVA	Analysis of variance
BAI	Beck Anxiety Inventory (a widely used research measure of anxiety)
BDI	Beck Depression Inventory (a widely used research measure of depression)
BME	Black and Minority Ethnic
BtB	Beating the Blues (computerised CBT therapy program)
CAB	Citizens' Advice Bureau
CBT	Cognitive Behaviour Therapy
cCBT	Computerised CBT
CI	Confidence interval
CM	Case Manager
CMHPs	Common Mental Health Problems (usually understood to include depression and anxiety)
CORE-OM	Clinical Outcomes in Routine Evaluation (34 item general mental health scale)
CSQ-8	Client Satisfaction Questionnaire (eight item scale)
CT	Cognitive Therapy
DH/DoH	Department of Health
EQ-5D	EuroQol Group 5-Dimension Self-Report Questionnaire
ES	Effect size (a way of expressing the difference between two distributions, eg of a measure given before and after treatment. An ES of 1.0 implies a change of 1sd between occasions)
GAD-7	Generalised Anxiety Disorder Assessment (seven item scale)

GP	General Practitioner
GSH	Guided self-help
GVA	Gross Value Added (previously described as GDP, Gross Domestic Product)
IAPT	Improving Access to Psychological Therapies
ICD-10	World Health Organisation International Classification of Diseases (tenth revision)
ICER	Incremental Cost-Effectiveness Ratio
IP	In-patient
IPA	Interpretative Phenomenological Analysis (method of qualitative data analysis)
IT	Information Technology
ITT	Intention to treat analysis (an analysis carried out including all patients referred to a trial or service, rather than just those who completed treatment)
LOCF	Last observation carried forward (the practice of using the latest of a series of observations in place of a missing planned, final observation)
MDS	Minimum Data Set
MH	Mental Health
MIQUEST	Morbidity Information Query Export Syntax (DoH sponsored data interrogation software for accessing computer held health records)
NHS	National Health Service
NIHR	National Institute for Health Research
NICE	National Institute for Health and Clinical Excellence
OCD	Obsessive Compulsive Disorder
OD	Organisational Development
OP	Out-patient
PBE	Practice based evidence

PC-MIS	University of York Patient Case Management Information System
PCT	Primary Care Trust
PHQ-9	Patient Health Questionnaire (nine item depression scale)
PIAG	Patient Information Advisory Group (governance group reviewing requests to access personal patient data)
PID	Project Initiation Document
PMS	Personal Medical Services
PPR	Primary percentage recovery rate (percentage of 'cases' falling below the agreed threshold for 'caseness' after treatment)
PWP	Psychological Wellbeing Practitioner (later name for IAPT staff initially called Low-Intensity Practitioners)
RCI	Reliable Change Index (the minimum amount of change on an outcome measure required to exclude measurement error)
SAPREL	Secure and Private Record Linkage (data linking software to permit anonymised data from two or more sources to be linked)
SD/sd	Standard deviation
SDO	NIHR Service Delivery and Organisation programme
SPSS	Statistical Package for Social Sciences (quantitative data analysis software)
STATA	A statistical analysis and graph drawing program (of particular use in carrying out and displaying meta-analyses)
SUS	Secondary care Use of Services (computer records of NHS service use)
wte	Whole time equivalent

Acknowledgements

This research could not have been undertaken and this report could not have been written without the help and co-operation of many people.

We wish to acknowledge the work of Abby Constantine and Sue Ridgway who provided excellent administrative and clerical support; Kate Doran, who undertook interviews and analysis for Chapter 4, and whose PhD research on user experience of 'Beating the Blues' informed this chapter; Rachel Horn, whose DClinPsych dissertation on the experience of 'stepping up' also informed Chapter 4.

Many other colleagues from Sheffield made contributions in various ways: Anna Thake (data inputting, CSQ analysis, work on chapter 4); Aimee Card & Saiqa Shah (matching cohort with download data, retrieval of papers for review; Mike Bradburn (consultation on Stata); Nik Vlissides, Laura Di Bono and Thom Kipling (data inputting/help with data preparation).

We should also like to thank those staff in Doncaster and Newham who very graciously extended their full co-operation with this research in helping us download data, obtain documentation, have access to patients and by providing comments on the draft report; these include Dr Rupert Suckling, Dawn White, Carole Hirst, Felix Davies, Mark Boon; Dr Ben Wright, Neelam Dosanjh, Julia Doe, Dr Anu Sinha.

We are grateful to the GPs and the Practice Managers in the practices which sent out questionnaires and allowed access to their IT systems for data downloads:

In Doncaster: The Health Centre (Dr Agrawal); Bentley Surgery; Auckley Surgery (Dr Okelele); Conisbrough Group Practice; The Lakeside Practice, Askern; The Vermuyden Centre, Thorne; The Sandringham Practice, Intake; West End Clinic, Rossington; St John's Group Practice, Balby; Dr Islam, Stainforth; Dr Jackson, Dunsville; Oakwood Surgery, Cantley; Field Road Surgery, Stainforth; Parkview Surgery, Sprotborough; The Scott Practice, Balby and Sprotborough; Princess Medical Centre; Church View Surgery, Denaby Main; Kingthorne Group Practice.

In Barnsley: Drs Kadar Sha & Muthulingaswamy; Drs Leese & Cornell, Thurnscoe Centre; Dr Khan, Galtee More Surgery; Walker Sykes PMS Practice /

Project 08/1610/154

Dove Valley Practice; Penistone Group Practice; Dr A N Dutta's Practice, Thurnscoe Centre; Goldthorpe First PMS Practice.

In Wakefield: Prospect Road Surgery, Ossett; Riverside Medical Centre, Castleford; Queen Street Surgery, Normanton.

In Newham: E12 Health Centre, Manor Park; Dr Shanker & Partners, Upton Lane; Glen Road Medical Centre, Plaistow; Dr Driver & partners, Lord Lister Health Centre; Dr Anu Sinha; Abbey Road Surgery, Stratford; Royal Docks Medical Practice

City & Hackney: Statham Grove Surgery, Stoke Newington; The De Beauvoir Surgery; The Hoxton Surgery; Lower Clapton Group Practice; Shoreditch Park Surgery

We appreciate the comments of members of the IAPT Management Board, including James Seward, Professor David Clark, Professor David Richards, Dr Alan Cohen and Jeremy Clarke, at various stages of planning and implementing this report, and for their comments on an earlier draft.

We also received very helpful comments on the draft from a panel of experts by experience including Dr Thomas Kabir (Institute of Psychiatry) and Stephanie de la Hay (Survivors of Depression in Transition – SoDiT).

Finally, we would like to thank all the hundreds of participants who completed numerous questionnaires and/or submitted to lengthy interviews to give us the data without which we could have done nothing.

Executive summary

Aim

To conduct a multi-facetted, 'whole system' evaluation of two services demonstrating a new model Improving Access to Psychotherapy Services (IAPT) for people with common mental health problems.

Research questions

- The descriptive question: What form do the services take, how are they organised, whom do they see, what is the patient pathway and how many people are seen?
- The patient experience question: What is the experience of patients using the system and how satisfied are they with it? What lessons can be learned from their experience to improve services?
- The organisational question: What are the workforce implications of establishing an IAPT service, and what lessons on implementation and feasibility can be learned from the demonstration sites?
- The efficacy question: What improvements in mental health and functioning seem to follow from the implementation of an IAPT service? How do these improvements compare to those achieved historically, to those reported in randomised trials and other investigations, and to those achieved in comparable NHS mental health services?
- The effectiveness question: Do the new services perform any better than existing services in matched locations for the same client group? If so, is the improvement cost effective (or conversely, can the services achieve comparable outcomes at lower unit cost?)

Methods

A range of quantitative and qualitative methods was used to address these evaluation questions:

To investigate *service delivery and organisation*, extensive documentary evidence from each service was collected together with data on referrals, assessment, treatment and clinical outcome, routinely collected by the two

services from June 2006 to April 2009. These data were anonymised and descriptive statistics were used to provide information on service inputs, processes, outputs and outcomes.

To investigate *organisational processes in implementing the innovation*, we conducted an organisational case study, interviewing a purposive sample of 57 stakeholders within the IAPT sites and analysing transcripts thematically.

To investigate the *patient experiences of, and care pathways through, IAPT services* we used a qualitative exploration of key themes using in-depth interviews with 77 patients and analyses informed by the Framework method.

To investigate *service effectiveness*, outcome data were analysed statistically and compared with benchmarks calculated from archived primary care mental health services datasets and from randomised controlled clinical trials in depression and anxiety disorders.

To investigate *cost effectiveness*, a postal questionnaire study of 504 people examined service costs and outcomes for primary care patients eligible for IAPT within the demonstration sites, compared with similar patients in comparison sites, matched on a range of socio-demographic factors.

To investigate *patients' access to IAPT and the impact of IAPT on use of hospital services, sickness certification and psychotropic medication*, we used an innovative health informatics method linking de-identified data from General Practice IT systems, secondary service use datasets and the IAPT datasets.

Main findings

The two services differed widely in their local context and the model of service they aspired to provide. Both demonstration services succeeded in improving access to psychological therapies. One site achieved a high volume capacity through providing collaborative care: primarily CBT-informed guided self help rather than formal cognitive behaviour therapy; the other evolved a model of delivery which balanced guided self help with a higher volume CBT service.

Other successful aspects of the demonstration services included the use of non-traditional access pathways such as self-referral and referrals through the 'pathways to work' programme. Also demonstrated were the use of community outreach and intensive efforts to engage hard-to-reach populations such as people from black and minority ethnic groups.

Difficult challenges for the services included the extraordinary time pressures to establish complex services and difficulties establishing genuine partnership working between primary and secondary care, with non-NHS organisations and with service users.

Both services were characterised by short waiting times for first contact, an assessment process that 'signposted' people to other services and referrals that included a wide range of severity of problem. The services were used by unemployed people and people on benefits, who tended to have more severe difficulties.

The organisational process of implementation provided valuable insights into helpful and hindering factors in implementation. Ways to improve partnership working were outlined, and the defining characteristics of the IAPT innovation emerged clearly in terms of addressing mental health inequalities, breaking down barriers and creating a service that supported individuals in remaining active in society and in their community. Moving away from traditional clinical delivery methods was cited as a key way of avoiding some of the stigma attached to mental health issues, overcoming shortfalls in support for those in employment who are finding it difficult to cope and in providing access to those in previously hard to reach communities i.e. those from black and ethnic minority communities and non English speakers. Achieving a genuinely seamless pathway by good collaboration between primary and secondary services was an enduring difficulty and requires continuing attention.

Discussing with patients their experience of the IAPT service showed the importance of the first contact with their GP and with the IAPT service in helping to identify the problem, provide hope and a way forward. This was particularly helpful when people had a sense of control and choice and were seen quickly. Self-referral was often associated with feeling greater self-confidence and hope. However, some patients experienced little or no choice in either referral or treatment options and information that could have helped in decision-making was often not available.

The best experience for patients in terms of guided self-help interventions was characterised by good communication and working with responsive, flexible, and respectful psychological wellbeing practitioners in a structured format tailored to their needs. Negative experience was reported when the practitioner was seen as impersonal, self-help booklets were not pitched at the right level and although there were patients who liked the freedom of telephone contacts and the computerized packages, many found them

problematic. Careful introduction, some one-to-one sessions and real support helped improve the value of telephone/computer working.

Cognitive behaviour therapy was generally valued, but was often thought to be too short. A lack of continuity (due to staff turnover) or follow-up was problematic for patients. Service users' testimony underlined the importance of genuine choice and consent rather than assuming consent from passive acquiescence in this model of service delivery.

In terms of outcomes, in terms of the least stringent criterion, both services fell only marginally short of the 50% recovery rate set by the Department of Health as the target for those receiving a minimum treatment of two or more contacts. The target rate of recovery was exceeded when considering those patients who completed their individually agreed treatment plans.

Comparisons with previous primary care therapy service outcomes suggest that the new services delivered a service of equivalent effectiveness despite being newly-established and delivered by relatively inexperienced practitioners; this is a considerable achievement. Comparison with results reported in research trials showed therapy effects were slightly less favourable, a finding consistent with previous research.

The postal questionnaire survey generally found patients' well being and mental health had improved over four and eight month intervals, but this was matched by improvements in the comparison sites, giving very little difference in outcome between the IAPT demonstration sites and the comparators, although poor response rates to the questionnaire throws doubt on the robustness of this finding. Resource use did not change significantly for most of the mental health services that were reported across the IAPT sites and their matched comparator sites, but GP consultations and other health service use in Doncaster reduced more than in the comparison sites.

This study compared the costs and outcomes of patients recruited the two demonstration sites of Doncaster and Newham with those achieved at comparator sites in Wakefield and Barnsley and City and Hackney respectively. Service costs were found to be £263 (95% CI: -£258 to £779) higher in Doncaster compared to Wakefield and Barnsley and £561 (95% CI: -£333 to £1,451) higher in Newham compared to City and Hackney over 8 months for IAPT. These additional costs of IAPT generated 0.007 (95% CI: -0.006 to 0.021) additional Quality Adjusted Life Years (QALYS) in Doncaster but was associated with QALY losses in Newham, -0.002 (95% CI: -0.035 to 0.031), compared to their respective matched sites. This resulted in an incremental cost effectiveness ratio (ICER) of £37,571 per QALY when Doncaster was

compared to Wakefield and Barnsley but the probability that IAPT was cost effective was below 40%. However, using QALY estimates based on EQ-5D predictions brought down the ICER to £20,230 but with a probability that IAPT was cost effective of just over 40%. Lost employment costs were higher for Doncaster compared to Wakefield & Barnsley, £279 (95% CI: -£65 to £624) but lower for Newham compared to City & Hackney, -£212 (95% CI: -£522 to £98) although these differences fell to £67 in Doncaster compared to Wakefield and Barnsley when outliers were removed. These results indicate that the Doncaster IAPT demonstration site provided a service that was probably cost-effective within the usual NICE threshold range of £20,000-30,000, but there was considerable uncertainty surrounding the costs and outcome differences and it was somewhat undermined by the low response rate to the patient questionnaire (though comparisons with the IAPT suggest this may have resulted in an underestimate of the cost effectiveness of this service). It is not possible to comment on the cost effectiveness of the Newham service since the numbers were too low and the comparator site adopted an IAPT service during the study.

The general practice study findings suggest that IAPT referral is being appropriately targeted on people with a greater severity of problem, sickness certification and use of medication, and although it is not reducing antidepressant prescribing overall, it seems to be reducing sickness certification and may lead to a reduction in the use of Accident and Emergency attendances. These potential benefits at the practice level are diluted by the small proportion (6%) of people with common mental health problems who are referred.

Conclusions

Results from the demonstration sites show that both services were successfully established and offered good access to collaborative care for people with common mental health problems. Results met Department of Health expectations and were equivalent to psychological therapies delivered by other primary care practitioners, with evidence of reduction in sickness certification and possibly in the use of some secondary health services. Return to work was demonstrated but not specifically attributable to the IAPT intervention. We were able to estimate incremental cost effectiveness for the Doncaster service, which gave a probable ratio within the range of the NICE threshold.

The Report

1 Introduction

This report describes an evaluation of the first two demonstration sites for the Improving Access to Psychological Therapies (IAPT) programme. The sites, in Doncaster (South Yorkshire) and Newham (East London) were set up by the Department of Health (DoH) and operated in their demonstration role from the summer of 2006. Their aim was to test a new way of making well-researched psychological therapies available to a wider public than had been able to receive them before. The plan was that common mental health problems such as mild to moderate anxiety and depression should be treated more quickly and effectively. An important part of this plan was to make available treatments specifically recommended by the National Institute for Health and Clinical Excellence (NICE). It was hoped that this would reduce the burden of unhappiness caused by mental ill health for a significant proportion of the population. Additionally, the idea was to help people to return to (or stay in) employment who might otherwise have been prevented from doing so by readily treatable conditions.

It is important to note that this evaluation is only of the two demonstration sites and not of the national IAPT scheme as a whole which has sprung rapidly from them. In some ways the demonstration sites were exceptional and are not typical of later, more general IAPT services. The very fact of being first in line meant that a greater resource was made available, and that equally a greater burden of scrutiny was placed upon them. Nevertheless the evaluation team hopes that the results of this evaluation will be useful to those planning, commissioning, running and using IAPT and similar services in the future.

Genesis of the IAPT idea

There are many places where the story of IAPT might be said to have started. Perhaps the most convincing time and place which can be identified is a discussion at the 10 Downing Street Strategy Unit which was held on 20th January 2005. At that meeting Lord Richard Layard presented a paper entitled *Mental Health: Britain's Biggest Social Problem?[1]* In the paper he proposed that if unemployment had been the biggest source of misery in the 1990s, that role had been taken on in the subsequent decade by mental ill health. He

argued that successes in tackling unemployment needed to be followed up by aggressive measures to tackle common, disabling mental health problems. He drew attention to the national Psychiatric Morbidity Survey, which suggests that at any one time 16% of adults suffer mental ill health, with knock on effects for their families and carers. He asserted that effective treatments, although developed, were not readily available because services did not exist to deliver them on the scale needed. He noted that, even when evidence-based psychological therapies were available, there were long waiting lists (6 months or more) before they could be accessed.

"Mental illness matters because it causes massive suffering to patients and their families, because it prevents them contributing fully to society, and because it imposes heavy costs on taxpayers."[1]

Lord Layard suggested that the cost of mental ill health went beyond the subjective and hard to quantify personal cost borne by sufferers. There were quantifiable (and very large) costs to the public purse as well as more widely to society as a whole – for example in lost production due to inability to work. He proposed that the total annual loss from all sources due to mental ill health could be valued at around £46 billion.

1.1.1 Development of the demonstration sites and the evaluation

It was agreed to run a small number of services along the lines suggested by Layard, to demonstrate whether and to what extent his ideas were borne out in practice. Over the next twelve months or so plans were laid for two demonstration sites, one in the north of England and one in the south. Some of the detailed descriptions of this early commissioning period are reported more fully in Chapter 3.

By February 2006 the two demonstration sites were being identified and the initial form of the services sketched out. The NHS Service Delivery and Organisation R&D programme (SDO) put out a call for an independent evaluation for the demonstration sites. The SDO call listed a number of findings which had led to the development of the IAPT programme. The evidence suggested that:

- People with depression and anxiety did not have access to an appropriate response in primary care

- Cognitive Behaviour Therapy (CBT), combined therapies (e.g. CBT and medication together), bibliotherapy and other 'talking therapies' were effective both for anxiety disorders and for depressive disorders[2]

- Patients expressed preferences for 'talking therapies' i.e. psychological therapies (DoH 2003)

- Delays of twelve months or more were common in accessing secondary care based psychological therapy services, which, often in the absence of appropriate service levels in primary care, were used inappropriately.

The NICE Guidance[3] was invoked to emphasise the need for and efficacy of stepped care with the implication that most of the "front-line" access to psychological therapies should be in primary care. Stepped care is a model which seeks to treat service users at the lowest appropriate service tier in the first instance, only 'stepping up' to intensive/specialist services as clinically required.

Thus the level of professional input is augmented gradually, until satisfactory health status is achieved. This is said to offer clinical and financial advantages that can benefit service users, service providers and commissioners.

At this stage the plan was for the two demonstration sites to bring together a model of multi-disciplinary delivery of psychological therapies for people with mild to moderate depression. The model was expected to include the following characteristics:

- A team approach to delivering therapies in a stepped care context.

- A hub and spoke model with outreach into primary care practices. The area to be covered by the team would be a Borough

- Therapy according to NICE guidelines [4,5], with appropriate follow up and medication if needed in addition to CBT, which is the main therapeutic intervention. The NICE 2002 and 2004 depression guidelines recommend

  - Step 1 watchful waiting

  - Step 2 self help including bibliotherapy, computerised CBT, and/or practice based counselling

  - Step 3 CBT and/or medication

- Strong leadership by a psychologist

- Best practice in terms of training, supervision and peer support for therapists of all professions who provide talking therapy, in particular those who provide primary care, e.g. counsellors, practice staff and general practitioners

- Access to employment and housing advice at team level

- Collection of data routinely on process and outcomes (the latter using validated outcome measures).

A short-term 'internal' evaluation of early results from the sites was led by Professor David Clark and Professor Lord Richard Layard in 2006-7. The NIHR Service Delivery and Organisation R&D programme sought tenders to commission an independent, longer-term research evaluation in 2006. A team drawn from the Universities of Sheffield, Manchester and Southampton was awarded the contract to undertake the independent evaluation. A separate subcontract for the extraction of anonymised primary care data from the two sites was awarded to St George's, University of London. (The results of that collaboration and nested study are reported here as Chapter Seven).

The two clinical services were commissioned during spring of 2006 and began accepting their first patients in July of that year. The contract for this evaluation was negotiated over the autumn and was in place by 1st December 2006.

Plan for the evaluation

The SDO evaluation brief called for a range of qualitative and quantitative measures of both process and outcome. The questions to be answered included:

- The descriptive question: What form do the services take, how are they organised, whom do they see, what is the patient pathway and how many people are seen?

- The patient experience question: What is the experience of patients using the system and how satisfied are they with it? What lessons can be learned from their experience to improve services?

- The organisational question: What are the workforce implications of establishing an IAPT service, and what lessons on implementation and feasibility can be learned from the demonstration sites?

- The individual efficacy question: What improvements in mental health and functioning seem to follow for patients from the implementation of an IAPT service? How do these improvements compare to those achieved historically, to those reported in randomised trials and other investigations, and to those achieved in comparable NHS mental health services?

- The access and population effects question: Do the new services perform any better than existing services in matched locations for the same client group? If so, is the improvement cost effective (or conversely, can the services achieve comparable outcomes at lower unit cost?)

1.1.2 Data to be gathered to answer the research questions

A range of datasets were available or could be collected to answer the above questions. There were some readily available data which could be accessed including the following:

- Service download data: Routinely collected process and outcome data gathered by the two demonstration sites as part of their clinical activity.

- GP practice data: Routinely collected information on practice-wide populations held by GP practices in the demonstration sites' PCTs.

- Documentary data: Evidence about the planning, development and operation of the service derived from internally generated reports, minutes etc.

- Benchmarking data: Reports of trials and other investigations in the peer-reviewed literature, as well as large anonymised databases of outcome data in NHS and other settings.

In addition to these existing data sources to be interrogated, new sources of data were planned:

- Service user qualitative data: Interviews with service users to explore their experience of the service, as well as written material such as notes made on questionnaires.

- Organisational qualitative data: Interviews with planners, commissioners, managers and clinicians involved in the demonstration sites.

- Comparison cohort data: Process and outcome data gathered from specially recruited cohorts in the demonstration sites and in comparable PCTs without an IAPT service.

The data sources and the questions to be answered cross cut in the following ways (see Table 1).

Table 1. Research questions and data sources

Research question	Data source						
	Service download data	GP practice data	Documentary data	Benchmarking data	User qualitative data	Organisational qualitative data	Comparison cohort data
Service description	✓		✓			✓	
Patient experience	✓				✓		✓
Organisational impact		✓	✓			✓	
Individual efficacy	✓	✓		✓			✓
Access and population effects							✓

As can be seen, most questions are answered by data from at least two sources, and most sources contribute to answering more than one question.

2 Description of the demonstration services in Doncaster and Newham

Introduction

This chapter addresses the descriptive question: What form do the services take, how are they organised, whom do they see, what is the patient pathway and how many people are seen?

It describes the geographical and service context for the demonstration services, briefly outlines the way they were commissioned and set up, and describes some key events in their development. It then provides a quantitative description of the process of patients' progress through the system, in terms of
- the number and types of referrals received,
- the number of patients contacted,
- the numbers of patients who received an intervention, of what type
- the proportion of patients who do not engage with the service, or who drop out
- waiting times, e.g. between referral and first contact.

These data give insights into key issues of service quality, including service accessibility, equity, coverage, acceptability and efficiency.

This chapter does not examine service outcomes (presented in Chapter 5) or organisational systems (described in Chapter 3).

Although the services in Doncaster and Newham both aimed to improve access to psychological therapies, they served very different populations and were grounded in very different local organisational contexts. Because their aspirations and service models differed accordingly, our evaluation design did not seek to make direct evaluative comparisons between the two sites, but rather to evaluate each within its own context.

In undertaking this analysis, we are grateful to both services, which co-operated fully in providing us with access to documents and to data downloaded from their IT systems. Each service had a huge volume of material documenting the initiation and development of their services; in Newham electronically stored documents were available, in Doncaster we had copies of all 'hard copy' documents.

The planned audit period was three years from 1st May 2006 to 31st April 2009, but no referrals were received before 1st June 2006. The demonstration period for IAPT ended in April 2008 and we followed up for one year to include the period of moving from centrally commissioned demonstration site to routine NHS service under local commissioning arrangements, in line with our evaluation of service sustainability.

The data collection system in Doncaster was designed by the University of York team supporting the project, just in time for the August 2006 start date, but then changed during the data capture period. Different variables and codes were used prior to developing the data collection system to incorporate the final IAPT minimum data set (MDS) requirements and, at the time of the download (5/5/09), there had been only limited updating of existing data. Correspondence with the Service and PC-MIS[1] suggested the data collection system was functioning well by November 2006 but the IAPT MDS was not fully included until September 2008. Therefore data from 1/9/08 – 31/4/09 may be considered the most reliable. Some key analyses have been replicated in this data subset in order to test robustness of findings where data quality could have affected results.

The computerised data management system in Newham took longer to implement, having been commissioned *de novo* from an independent IT company.   This extremely arduous undertaking was reflected in additional project costs both direct and indirect service costs. Initially paper records were kept and then entered into the IT system retrospectively when it was fully functioning. Service redesign in Newham in 2007 necessitated some re-specification of the system, which was in operation by August 2007.   The key focus of the service in terms of data quality was demographics, clinical outcome data, contact time and final position on care pathway which indicated whether patients had received a service. The resulting IT system has since become fully functional and has been adopted by other IAPT services.

Doncaster

Doncaster is a large town in South Yorkshire; the resident population of the Metropolitan District in June 2008 was 289,300. Economically, it has a midsized local economy with a below-average GVA[2] per head. Unemployment

---

[1] PC-MIS is the University of York Patient Case Management Information System for the IAPT programme
[2] 'Gross Value Added' (GVA) is a measure of the value of goods and services produced in an area of an economy., now used to denote estimates that were termed Gross Domestic Product (GDP).

is above average (6.1% compared with 5.6% in Yorkshire and the Humber region and 5.4% in England) with a high claimant count (19% compared with 15% for Yorkshire and the Humber and 14% in England) and rates of pay are below average for full and part time workers. Employment in the area is dependent on Public Administration, Education and Health with high proportions in Transport and Communications and low proportions in Banking, Finance and Insurance relative to the regional average. Businesses in Doncaster are similar in size to the Yorkshire and Humber average with the vast majority employing fewer than 11 people although most of the labour force work for companies which employ 200+. People in this local authority have lower levels of educational qualification than the regional average with almost a sixth of the working age population having no qualification.

2.1.1 Doncaster service model and vision

The Doncaster service model pre-dated IAPT; the West PCT had undertaken a 'whole system review' of clinical services and implemented the stepped care model, using graduate workers to deliver facilitated self-help.   Doncaster PCTs already had strategic partnerships with Chamber of Commerce and the idea that NHS could work with employers and DWP to tackle problems of worklessness was well established. Prior liaison with Professor Richards at the University of York gave the foundation for this vision, which was an innovative implementation of stepped care, with IAPT case managers – not CBT experts – providing fast, open access to 'low intensity' support and information in primary care, for everyone with depression and anxiety. The service model was grounded in the value of 'partnership' working between health and other agencies, including employers, Chamber of Commerce, occupational health services and DWP initiatives (Pathways to Work, Condition Management Programmes).   Doncaster West PCT worked with the secondary mental health care Trust in putting the bid together. In October 2006 the PCTs in Doncaster merged into one Doncaster PCT and the service was commissioned Doncaster-wide.

The distinctive aspiration was a 'whole system approach' between partners working in a 'virtual organisation', with clinical, managerial and service user leadership, common IT systems to support delivery and evaluation, case managers supported by CBT, plurality of providers, next day appointments and an emphasis on choice including the community/voluntary sector and independent sector. "No single organisation can meet the needs we have identified, including the delivery of psychological therapies – one of the strongest messages from the Doncaster research is that each party has something to gain and *all* have a role to play" (Ref: Barriers to Employment

Working Group.) A 'hub' was set up in premises belonging to the Chamber of Commerce, acting as the resource centre and administrative base for the case managers, the channel through which referrals would be received. The aim was for the hub to operate outside normal working hours to help employed people and the original intention was to facilitate a wide source of referrals including self-referral, and to act as 'brokers' for referral to community resources, liaising with workplaces, voluntary and non-statutory services and job centre plus advisors. The original Project Initiation Document (PID) was focussed on benefit claimants; understanding their journey and how psychological therapies can help.

2.1.2 Doncaster service set up and structure

When central funding was agreed for two national demonstration sites, Doncaster was invited to tender, and a bid was submitted at the end of November 2005. The bid met eligibility criteria and local governance arrangements were put in place in January 2006 although funding was not formally confirmed until the end of April 2006. The managerial structure included
- Project Board
- Implementation Group
- Project director
- Team Manager
- Clinical Director of Psychological Therapies

A tendering process for an external Trainer and organisational development (OD) consultancy was undertaken and awarded to the University of York in April 2006 and the first case managers were appointed to work at the Hub in June 2006.   Staff addressed a number of challenges in setting up the service, including the difficulty of establishing suitable accommodation and ensuring NHS computer access from a non-NHS building. The IT system was an integral part of the training and OD package provided from the University of York, and combined a case supervision method, a patient management system and an outcomes data collection tool. People using the service were contacted by telephone and offered a face to face consultation, followed by telephone 'sessions'.   People were seen in GP surgeries and other community locations.

2.1.3 Doncaster service development

The service in Doncaster did not experience any major shifts of policy or practice, but rather had a developmental pattern of responding to operational and logistical difficulties as they arose. An early and rather persistent difficulty was the relative lack of cognitive behaviour therapists to provide a 'step 3'

service or case manager supervision. The original plan was that CBT practitioners in the secondary mental health trust could be deployed for this task, but by March 2007 it became clear that this was impractical because of their existing waiting lists.   Relatively few service users in Doncaster therefore accessed Step 3 interventions, partly by design and partly through necessity.

Doncaster service statistics

2.1.4 Referrals and waiting times

Over the 35 month period 1/6/06 – 30/4/09, 10,790 people were referred of whom we estimate 493 were re-referred at least once during the audit period[3], making 11,283 referrals in total, an average of 322 per month. The rate of referral increased steadily as the service became established, reducing slightly only for the last six months of the period. New referrals reduced in the last year.   New referrals and re-referrals for six month periods are shown in Figure 1.

Figure 1. Referrals in Doncaster by 6 month periods

---

[3] When someone was first seen in the service then re-referred later, the old record was overwritten so that data from the first referral was lost, and it was not possible to distinguish where one episode ended and another began. We therefore had to infer the number of re-referrals from internal evidence. For some analyses these are omitted as the data were not reliable. Where they are included, data relate to the most recent episode.

Patients referred to Doncaster Service
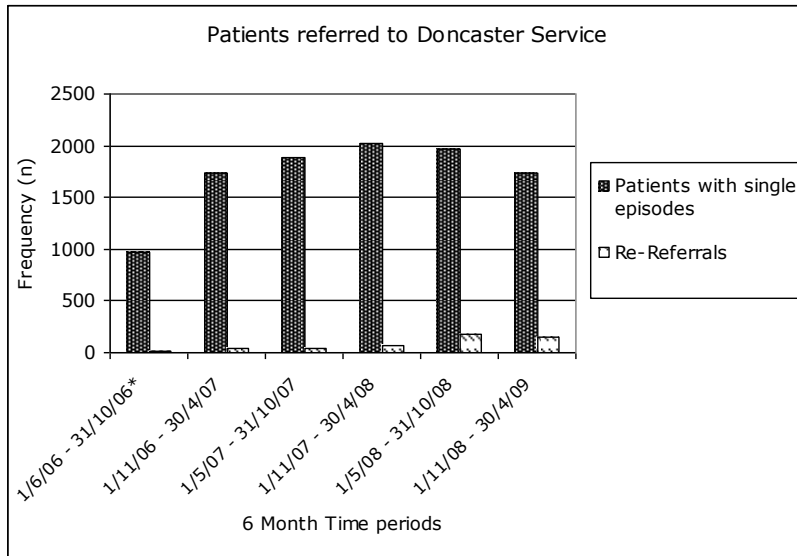
Figure 2 shows the flow of referrals through the service system, excluding re-referrals, i.e. for the 10,297 people with a single episode. Of these, 513 were still in treatment at the end of the audit period and 2932 had completed a course of treatment. Disregarding those still in treatment, those who completed were 30% of those referred, 46.9% of those contacted and 69% of those taken on for treatment.

Figure 2. Doncaster Referral Flow

Referred: 10297 Referred (R: 100%)	381 (3.7%) – Not suitable
	543 (5.3%) - Declined Treatment
	747 (7.3%) – Dropped out/never seen

2799 (27.2%) No Contacts

1 - Only age, gender & referral date - no other data
1122 (10.9%) – Completed Treatment (no contact data)

83 (0.8%) — No contact, reason unknown.

645 (6.2%) - Awaiting 1st Appointment

Contacts: 6770 (C: 100%; R: 65.8%): At least 1 contact (includes 14 with number of contacts unknown)	774 (11.4%) - Episode Complete
	280 (4.2%) – Not suitable
	217 (3.2%)- Declined Treatment
	462 (6.9%)– Dropped out
	3 (0%) - Deceased
	48 (0.7%) Completed one episode- (Ending type unknown)

1784 (26.3%) 1 Contact only

211 (3.1%) - Awaiting 2nd Appointment

Treated: 4761 (T: 100%; C: 70.3%; R:46.2%)	268 (5.6%) – Not suitable
	226 (4.7%)- Declined Treatment
	728 (15.3%) – Dropped out
	3 (0.1%) - Deceased

91 (1.9%) – Completed one episode- (Ending type unknown)

513 (10.8%) Currently In treatment

Completed treatment: 2932 (T: 61.6%; C: 43.3%; R: 28.5%)

R=referred; T=treated; C=completed treatment

Table 2.   Source of referrals[4]

SOURCE	N(%)
GP	9,712 (90.0)
Community Practice Nurse/Health Visitor	289 (2.7)
Other Clinical Speciality	263 (2.4)
Self referral	251 (2.3)
Job Centre Plus	107 (1.0)
Local Authority Social Services	62 (0.6)
Employer	2 (0)
Voluntary Sector	1 (0)
Education Service	1 (0)
Other	97 (0.9)
Missing	5 (0)
Total	10,790 (100)

Table 2 lists the source of referrals showing that 90% were from General Practitioners in Doncaster, with less than 5% from non-NHS sources.

---

[4] This table excludes re-referrals, but percentages would be unaffected by their inclusion.

Project 08/1610/154

2.1.5 Case mix

Table 3. Demographic profile of the referrals and the Doncaster population.

		IAPT patients	Doncaster population
Sex	% Men[5]	34.9	48.8
Age	Average age (s.d.)	38.2 (13.5)	
	% Over 55	11.6	27.3
Ethnicity	% Black/Asian[6]	0.4	2.1
Employment[7]	% employed	49.2 (51.1)	71.1
(7 mth subset)	% receiving benefits	40.3 (29.4)	21.1
	% on sick pay	21.1 (31.1)	

Table 3 shows that men, people over 55 and people from ethnic minorities did not access the service to the same extent as women, younger people and White people or as much as would be expected from the Doncaster population statistics. With regard to the gender mix, the proportion of men referred is still significantly higher than figures available for referrals to primary care counselling services (28.5% men; 95% CI 27.4-29.6)[2]. As would be expected, fewer people referred to the service were employed and more were claiming benefits than for the Doncaster population as a whole.

---

[5] Men were also less likely to be re-referred than women, at 28.4% of re-referrals

[6] This term encompasses Asian/Asian British, Black/Black British, Chinese, Mixed, and Other ethnic groups.

[7] For those with at least 1 contact with the Service (n=6770) 87.1% had no Employment Status recorded at first contact, 52.5% had no Sickness Pay data and 27.0% had no Benefits data recorded. We examined the 7 month data subset (n=1099) to check if data quality was better, yielding equivalent missing values of 33.7%, 29.7% and 32.6% respectively.

Table 4. Primary diagnosis assigned by referrer[8]

Diagnosis	ICD-10 codes	N	Valid %
Depressive disorder	F32,F33	7477	75.86
Mixed depression/anxiety	F41.2	1441	14.62
Anxiety/OCD/stress disorder	F40.0, F40.1, F40.2, F41.1, F42, F43.1	838	8.50
Other	F10, F31, F45, F50, F99	44	0.45
None	Z53.4	56	0.57
Missing		934	

Table 4 indicates three quarters of patients were referred with a depressive episode or with recurrent depressive episodes. This seems a high proportion, but the majority where a secondary diagnosis was recorded (93% of 5409) had F41.1, Generalised anxiety disorder. This suggests that most people referred were experiencing a combination of depression and anxiety.

The average duration of the patient's problem was 31.0 (s.d. 62.5) months with a range of 0 – 847 months.[9] Almost a quarter of people seen (for whom this was recorded) had mental health difficulties of less than six months duration (n=1622; 23.9%) and 518 (14.1%) of those seen had a problem duration of more than five years.

---

[8]Diagnoses were recorded by the service for only 916 (13.5%) of patients who had at least 1 contact. Referrer assigned diagnoses were therefore used.

[9] Duration of problem was calculated as the difference between the recorded date of onset and the referral date; 3668 (54.2%) of those seen had both dates recorded.

Project 08/1610/154

Severity of depression and anxiety was measured using the PHQ-9 and GAD-7, with mean (s.d) values of 15.80 (6.45) and 13.71 (5.35) respectively.[10] For depression, 81.2% of patients scored above the clinical cut-off on the PHQ (>=10) and 33% scored in the 'severe' range. The equivalent figures for anxiety on the GAD-7 (>=10) were 77.3% and 50.6%.

People from Black/Asian ethnic groups had significantly higher levels of depression and anxiety than those from White ethnic groups.[11] Those in receipt of benefits had markedly more severe depression and anxiety at referral.[12]   Women had slightly more severe depression and anxiety than men, which was statistically significant given the large sample.[13] A more severe level of anxiety and depression at referral was significantly associated with non-completion of the intervention.[14]

Data quality on medication was poor; information on medication at referral was not recorded for 87.6% of people with one or more contact, and 86.7% of data on medication at any point in the care spell were missing. For this reason, any analysis is likely to be misleading and unrepresentative of the whole sample, so is omitted here.

2.1.6 The service provided in Doncaster

Excluding re-referrals, 7028 (68.3%) had a referral date and first appointment date recorded, and for patients successfully contacted, the median waiting time was 3.6 weeks: 6.4% of patients had a first appointment date within one week of their referral, 22.4% within 2 weeks, 57.0% within 4 weeks and 76.0% within 6 weeks. The average and median waiting time gradually increased over the audit period (Figure 3)

---

[10] Severity is reported for patients where a single episode is indicated and who had at least 1 contact (n=6770)

[11] PHQ9: $t_{(6303)}$ = -2.44, p=0.015; GAD-7: $t_{(6292)}$ = -2.21, p=0.027.

[12] PHQ9: $t_{(4929)}$ = -17.04, p<0.001; GAD-7: $t_{(4924)}$ = -11.65, p<0.001.

[13] PHQ9: $t_{(6670)}$ = -2.16, p=0.031. GAD-7: $t_{(6659)}$ = -3.54, p<0.001.

[14] PHQ9: $t_{(5822)}$ = -8.32, p<0.001; GAD-7: $t_{(5814)}$ = - 6.20, p<0.001.

Figure 3. Waiting times for six month periods
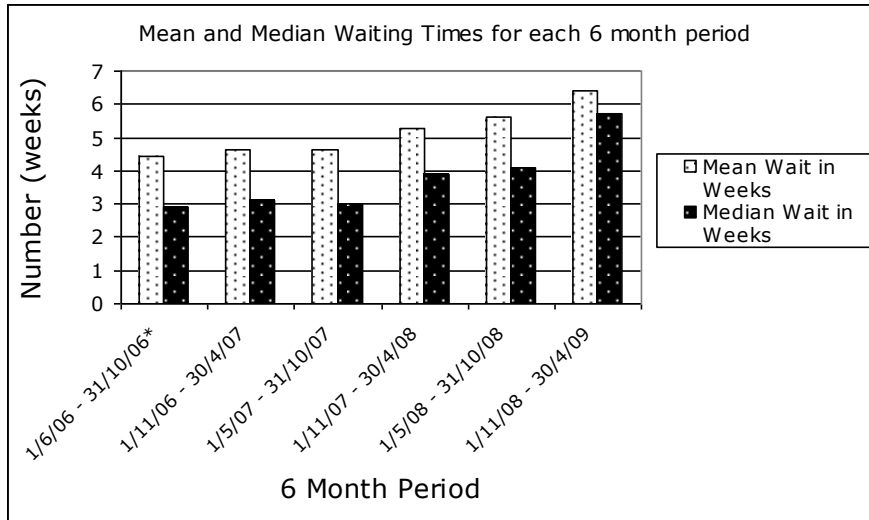
Information on the first assigned treatment step after contact with the service was available for 3562 (52.6%) of 6770 patients (Table 5).

Table 5.   Initial treatment step and baseline scores

First recorded Step	At least one Contact n(%)	Baseline PHQ-9		Baseline GAD-7	
		n (%)	Mean(sd)	n (%)	Mean(sd)
Step 1: Monitoring	155 (4.5)	149 (96.13)	13.23 (7.55)	149 (96.13)	11.68 (6.16)
Step 2: Guided self help	3226 (93.2)	3150 (97.64)	15.95 (6.36)	3142 (97.40)	13.83 (5.26)
Step 3: CBT	181 (5.2)	176 (97.24)	16.70 (7.13)	175 (96.69)	14.36 (5.79)

For people with at least one service contact, the great majority (93.2%) were initially allocated to a Step 2 treatment. Only 4.5% were initially assigned to Step 1 and 5.2% to Step 3. People assigned to step 1 had less severe depression and anxiety scores compared with those who enter at step 2, but there was no significant difference between the scores of those assigned to step 2 or step 3, so the basis on which some people go straight to step 3 is

less clear[15]. Of those with 'step' data, 180 (7.0%) were 'stepped-up', while 155 (6.1%) were 'stepped down' from their initial step during their treatment. Of those 'stepping-up', 156 went to CBT after having a step 2 intervention. These were found to have similar baseline scores to those who receive step 3 CBT from the outset (PHQ-9 mean (sd) 17.76 (6.08) and 16.57 (7.34) respectively; GAD-7 mean (sd) 14.94 (5.24) and 14.48 (5.98) respectively). The total number of people who received a CBT intervention was 337.

Excluding re-referrals, 4761 of the 10,297 referrals received an intervention, defined as two or more contacts with the service. These scored higher on the baseline PHQ-9 and GAD-7 measures than those patients who had only a single contact; 16.12 (6.19) compared with 15.01 (6.99) for the PHQ-9 ($t_{(6657)}$ = -6.39, p<0.001) and 13.98 (5.15) compared to 13.04 (5.77) for the GAD-7 ($t_{(6647)}$ = -6.50, p<0.001). Men and people under 40 were less likely to receive an intervention. A disproportionate number of men under 40 had only one contact, compared to women and men 40 or over ($\chi^2_{(1)}$= 14.99, p<0.001).The mean (sd) number of contacts, was 5.2 (4.0) with a range from 2 to 39. The modal intervention was two contacts, received by 1243 (26.1%) of patients. Three-quarters of the referrals had six contacts or fewer, only 5.3% of patients had more than 12 contacts. The majority (57.8%) of contacts were by telephone while 42.1% were face-to-face. A very small proportion (0.1%) was of another type, including text or email. The type of contact varied according to treatment step; step 3 (CBT) patients had more than twice the number face-to-face contacts than those on steps 1 or 2.

The main interventions offered at treatment step 2 were guided self-help, providing information, medication support and signposting to other services. Computerised CBT was used by only 295 (14.4%) of step 2 patients.

Between 1/6/06 and 30/4/09, 62 clinical staff were employed, of whom 53 saw patients at first contact; 21 of these provided step 3 ('high intensity') interventions and 41 provided step 2 ('low intensity') interventions. Over this period, 17 staff left the service; most of those who left (12) were trained 'low intensity' workers. Despite the 27.4% turnover, most people (89.5%) saw the same worker throughout their contact with the service.

---

[15] One-way ANOVA and post hoc analysis (Hochberg's GT2) indicated a significant difference for PHQ9 and GAD-7 baseline scores between Step 1 and 2 (mean differences 2.72, p<0.001 and 2.15, p<0.001 for PHQ9 and GAD-7 respectively) and between Steps 1 and 3 (mean differences 3.46, p<0.001 and 2.68, p<0.001), but not between Steps 2 and 3 (mean differences 0.75, p=0.353 and 0.53, p=0.414).

Newham

The London Borough of Newham is an area in the East End of London, five miles East of the City of London, north of the river Thames. It comprises West Ham, East Ham, Plaistow, Forest Gate, Manor Park, Beckton, Stratford, and Canning Town. In June 2008 it had a population of 249,500 of which 165,920 were of working age. Newham is characterised by ethnic diversity; 39% of the population is White (compared with 69% for London as a whole and 88% for England), 34% are Asian or Asian British, 20% are Black or Black British and 7% of other ethnicity.   Of the White population, 19% are White of non-British ethnicity such as people from Kurdish or Eastern European backgrounds. The population also has a relatively high proportion of men, of people under 45 years old, and of single people and one person households. In 2001 it was the highest ranking borough in England and Wales for lone parent households with dependent children.   It is an economically deprived area with relatively high rates of unemployment (11.3% compared with 6.9% for London and 5.4% for England) and benefits claimants (20% of people of working age compared with 14% for London and England).

2.1.7 Newham service model and vision

Pre-dating IAPT, statutory services had been providing a broad range of psychological therapies to the people of Newham over many years. There was a Borough wide service (serving all GPs) for people with mild to moderate mental health problems which offered a brief intervention service of up to eight appointments. Therapeutic interventions were provided by psychotherapists, psychologists and counsellors based in GP surgeries and referrals were received directly from GPs. It provided brief psychodynamic, CBT and systemic therapies and a range of counselling, plus self help materials and signposting to voluntary sector and community support resources. In October 2006 this consisted of 6.7 whole time equivalent (wte) staff. There was a generic 'tier 2' service, provided by 12 wte psychologists with extensive links and consultation to third sector providers, delivering Sure-Start peri-natal training and service and a transitional team for refugees and asylum seekers. In addition, since 2001, through PMS pilot funding after the 2001 DH Guideline on Treatment Choice in Psychological Therapies [6], a CBT service had been provided by nurse therapists at three locations in Newham; Wordsworth Health Centre, Essex Lodge and The Project Surgery [7].

The Newham service was therefore designed to complement and be integrated with these services for "the cohesive delivery of evidence-based psychological

therapies with greatly enhanced access and choice"[8]. The original aspiration was to provide cognitive behaviour therapy, of the type that had been found effective in randomised controlled trials and which formed the basis of NICE guidance. The innovation was that therapists would work across a range of community locations, of which GP surgeries where to be the initial targets, in order to promote ease of access and make the service more responsive as well as in a 'Psychological Treatment Centre' where the therapists would receive peer support, supervision, training and could participate in other activities associated with clinical governance and professional development. This space would also be used to deliver therapy and provide access to more specialist or expert therapies where required. In addition, culturally appropriate psychoeducational and self-help materials based on CBT would be developed and translated into local languages and large group CBT-based psychoeducation would be available. The proposal specifically addressed the needs of under-represented groups (such as the unemployed, people of Black and other Black/Asian ethnicity). This included a plan to foster self-referral, to employ bilingual therapists and to engage in active outreach within the Newham community. In promoting access to groups which were hard to reach, Newham IAPT had a very strong emphasis on engagement – with up to five phone calls and three letters to each referred person to facilitate engagement. Alongside this service, an employment Support Service would be provided by Mental Health Matters, a voluntary sector agency. Employment Coaches would assess, co-ordinate training, place and support service users in gaining and maintaining employment. The PID also emphasised the ways that the new Cognitive Behaviour Therapy Service would dovetail with other psychological therapy services available in the Borough. The final PID suggested that the funding available for the demonstration site would not support coverage of the whole Borough and that the service would in the first instance be available to 30% of general practices. The final PID also emphasised that service users would "be involved in all stages of the planning, delivery and evaluation of the Newham Pilot and local service user groups such as the Independent Newham Users Forum will be actively involved in a Steering Group, overseeing the delivery of the Pilot".

2.1.8 Newham service set up and structure

Newham was invited to become a national demonstration site in October 2005 and the programme plan was submitted at the end of November 2005. In January 2006 a Steering Group and an Operational and Clinical Group were set up and the Clinical Lead was in post. In the first phase of the service, it was available to 13 General Practices in Newham. Staff addressed significant challenges setting up the new service to a very demanding timeline, including

therapist recruitment, developing an IT system, funding uncertainties, providing early outcome evidence for the spending review and finding adequate accommodation.   From confirmation of funding in March 2006, to the first patients being seen was less than five months and the first data reports were made only three months later.

The new IAPT CBT service in Newham was the focus of this evaluation. In addition, a proportion of IAPT funding was provided to enhance existing provision, including from February 2007 some additional systemic therapy provision.   As part of integrating with other services, the assessment process therefore included consideration of whether an alternative therapeutic approach would be more appropriate and this 'signposting' function was an important aspect of the service.

2.1.9 Newham service developments

Funding for the project was confirmed in March 2006 and the Newham project constructed to run for an 18 month period therefore the initial funds were to cover this period. At the end of the financial year 2006-2007 funding for a further year was confirmed, however some staff were on a fixed term contract for the anticipated project duration which affected staff retention.   When funding was confirmed, the Department of Health required Newham to extend coverage to the whole Borough and to increase throughput. This led to phase 2 of the Newham service, where the original vision was modified to include a greater provision of 'step 2' interventions by employing more 'assistant therapists'[16].   In addition, a new pathway was implemented which used a brief structured telephone assessment ('telephone triage') where a qualified therapist determined whether low intensity care was or was not appropriate, with the option of a face to face assessment if needed. If low intensity work was contraindicated, the patient could be directed to high intensity CBT or another service.

The reason for this change was to reduce the negative experience of some patients who developed a strong therapeutic relationship in the face to face meetings who were then disappointed by their move to another therapist. The new arrangement also directed the clinically appropriate group to low intensity care.

The resulting model was intended to be a robust hybrid between the original Newham model and the Doncaster model.

---

[16] This term was used in Newham at this stage to refer to Grade 5 and 6 workers (otherwise termed 'case managers', 'low intensity workers' or more recently 'psychological wellbeing practitioners').

2.1.10   Newham service statistics

2.1.11   Newham Referrals and waiting times

In the 33 months between 23/6/2006 and 31/3/2009, 3349 people were referred to the service, of which 259 were referred more than once: 238 had two episodes, 20 had three episodes and 1 had 4 episodes recorded[17].   The majority (64%) of the 259 people who had more than one episode did not attend any intervention sessions at their first referral[18].

The operational policy of the service changed in 2007, when a high volume, low intensity service was implemented at the request of the Department of Health. The rate of referral did increase, from an average of 54 per month for the first three months, to 132 per month for the final six months. The change in care pathway and system structure resulted in an increase in patients whose first intervention session was low intensity from 7% in the first 6 months to 66% in the last 6 months of the evaluation period. There was also an increase in service throughput from under 10 patients per month in 2006 to over 80 patients per month in the first half of 2009. New referrals and re-referrals for six month periods are shown in Figure 4.

---

[17] This gives a total of 3371 referrals.

[18] Many re-referrals (35%) were technical re-referrals in that they are recorded as 'failed to engage' at their first referral and the case is then closed, but contact is subsequently made, often shortly after case closure.

Figure 4. Referrals in Newham by 6 month periods
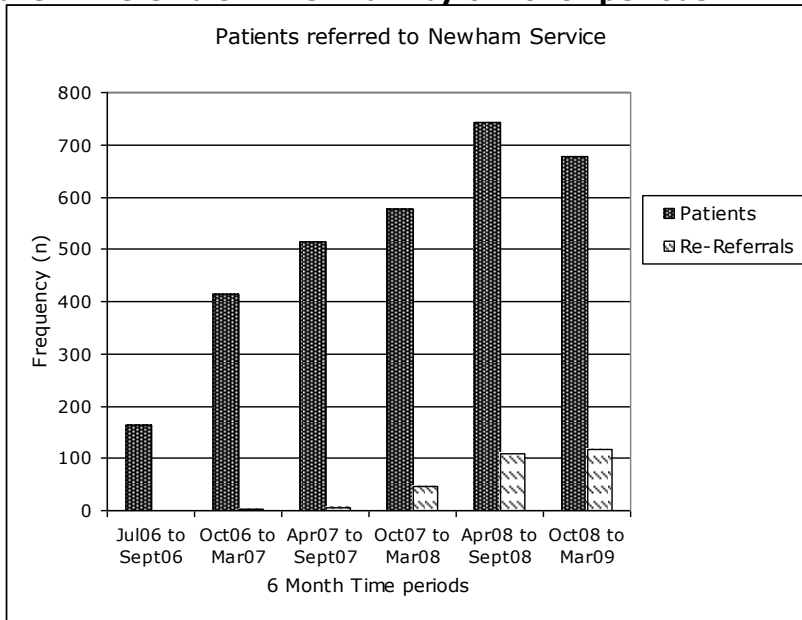
Figure 5 shows the flow of referrals through the service system. Of 3349 people referred, 320 were still in treatment at the end of the audit period and 592 had completed a course of treatment. Disregarding those still in treatment, those who completed were 19.5% of those referred, 24.5% of those contacted and 51.6% of those taken on for treatment.

Figure 5. Newham Referral Flow Chart
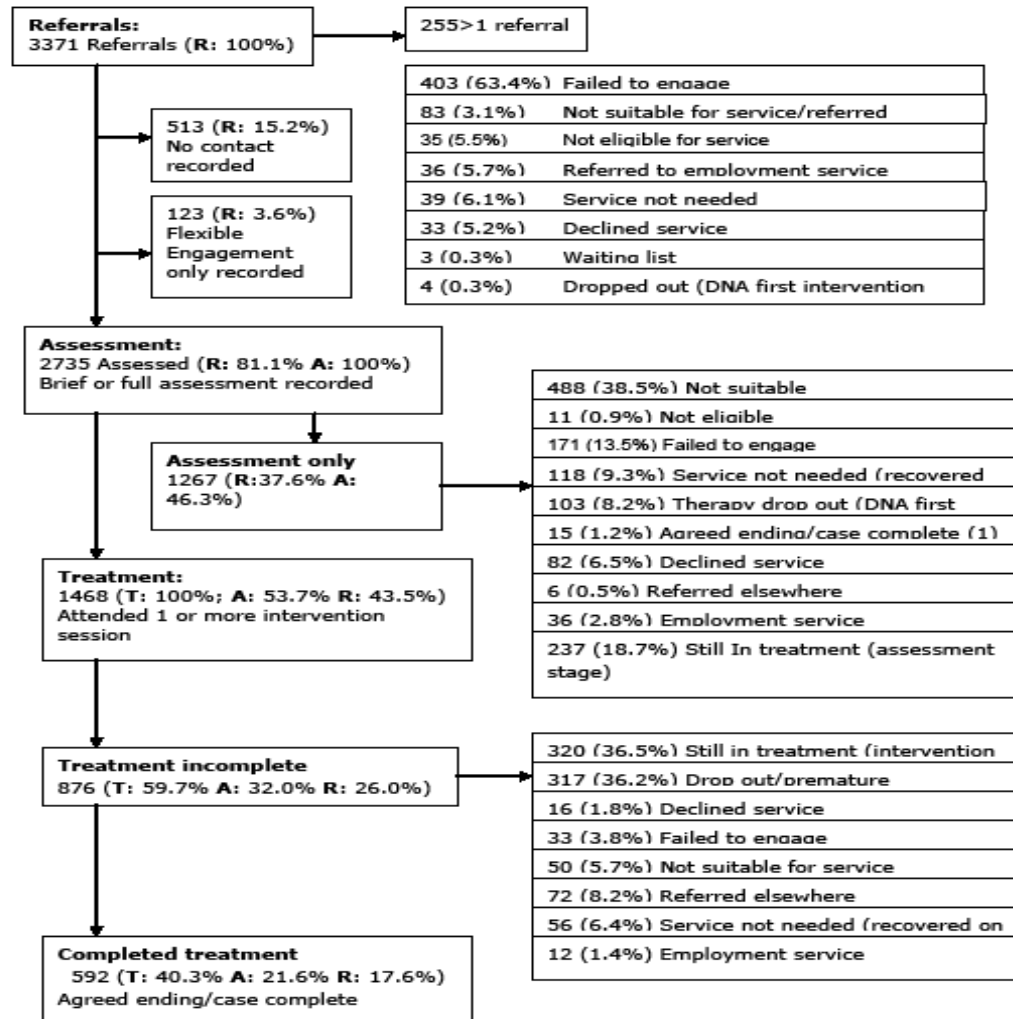
Table 6. Source of referrals

Source	N(%)
GP	2208 (65.6)
Other Professional	284 (8.4)
Pathways to work	112 (3.3)
Self Referral	764 (22.7)
Total	3368

Table 6 lists the source of referrals showing that 65.6% were from GPs but that a substantial proportion, almost a quarter, were self-referrals. This is a feature of the Newham service which was in line with its operational policy. To examine this more closely, we examined changes in the numbers and proportion of self-referrals over time, where it emerged that after a modest rate of self-referral in the first nine months of the service (averaging about 10%), there was a marked increase in the numbers of self referrals in the second year, which was sustained through the audit period, although the proportion of self referrals dropped because of the increase in the total number of referrals from other sources (see Figure 6).

Figure 6. Total referrals and self-referrals in six month periods

Total and self referrals in Newham
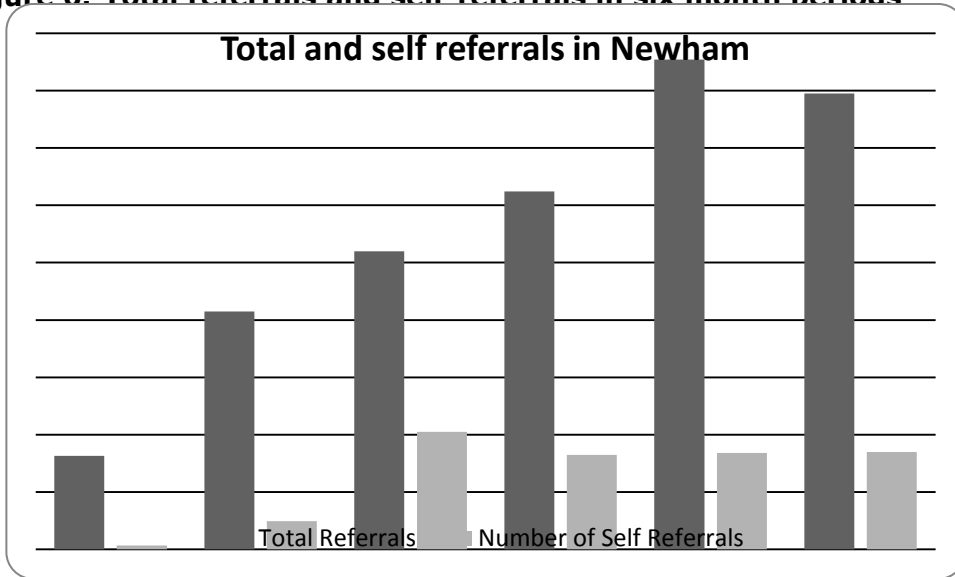
2.1.12 Newham Case mix

Table 7. Demographic profile of the referrals and the Newham population[19].

		Newham IAPT patients	Newham population[20]
Sex	% Men	38.9	50.9
Age	Mean (SD)	37.7 (12.3)	
	% Over 55	9.13	15.5[21]
Ethnicity	% Black/Asian	51.9	61.0
	% Asian/Asian British	27.8	33.9
	% Black/Black British	18.4	19.6
	% Chinese/mixed/other	5.8	7.5
Employment	% employed	43.2	58.4
	% receiving benefits	38.6	20
	% on incapacity benefit	17.5	8

Table 7 shows that men and people over 55 did not access the service to the same extent as women or younger people or as much as would be expected from the Newham population statistics. The proportion of men is significantly greater than figures available for referrals to primary care counselling services (28.5% men; 95% CI 27.4-29.6)[9]. There was no gender difference in the rate of re-referrals (39.1% male).

More than half the referrals were people from Black/Asian ethnic groups, although compared with the Newham population they were still slightly under-

---

[19] Demographics have been calculated for cases (n=3090) rather than referrals (n=3371). For people who were re-referred (n=259) data from the last episode has been used.

[20] Data from ONS Neighbourhood Statistics, data tables: Key Figures for People and Society; Population and Migration, Key Figures for Economic Deprivation; Age UV04.

[21] Proportion of the Newham population age 17-74 over 55.

represented, particularly the Asian/Asian British subgroup compared with the Black/Black British group. When the data are viewed longitudinally a steady increase in the rate of referrals from Black/Asian groups is observed. This is primarily a rise in referrals from the Black/Black British population. When people of White minority ethnic status are taken into account, there is overall no evidence of major inequity with regard to ethnicity.   Although the majority of people were referred via their GP regardless of ethnicity, compared with the White group, people of Black/Asian ethnicity were rather more likely to access the service via self-referral (25.7% vs 21.5%). In terms of language, 11.6 percent of all patients were recorded as not speaking English and 14.5 percent as not reading English.

Table 8.   Primary diagnosis assigned by service

Diagnosis	ICD-10 codes	N	Valid %
Depressive disorder	F32, F33, F34.1	807	42.9
Mixed depression/anxiety	F41.2	162	8.6
Anxiety/OCD/stress disorder	F40, F41, F42, F43, F45, F48	746	39.6
Other	F10, F20, F31, F60, R45.4	153	8.1
None	Z codes and physical injury	15	0.8
Missing[22]		852	

Table 8 shows that 43% of referrals which were assessed and assigned a primary diagnosis had a depressive disorder; of these, 16% were recurrent or persistent. There were a wide range of anxiety disorders; the most prevalent were generalised anxiety disorder (7.3% of all referrals), post-traumatic stress disorder (6.3%) and panic disorder (6%). Obsessive compulsive disorders constituted 2.9% of referrals.

Data on the duration of the current episode were available for 57.4% (n=1775) of people referred[23]. The majority (71.2%) had been experiencing problems for over a year, and 14.2 percent for over 10 years.

---

[22] Missing values are calculated as a proportion of those referrals which were assessed, n=2735.

The severity of depression and anxiety was measured with the PHQ-9 and GAD-7, with mean (s,d) values of 14.73 (6.27) and 12.5 (5.39) respectively. For depression, 78.2% of patients scored above the clinical cut-off on the PHQ-9 (>=10) and 25.6% scored in the 'severe' range. The equivalent figures for anxiety on the GAD-7 (>=10) were 72.9% and 44.3%.

Those in receipt of benefits had significantly more severe depression and anxiety at referral.[24] Women had significantly more severe depression and anxiety than men.[25] A more severe level of depression at referral was significantly associated with non-completion of the intervention[26], but initial severity of anxiety did not predict this.   There was no association with initial severity of symptoms for age, ethnicity or duration of current episode.

Data on medication were available for 50% of patients referred. Over half (54.6 percent) were taking medication, the majority of these (63.3 percent) SSRI anti-depressants, 9.6 percent other antidepressants, 6.2 percent sedatives and 18 percent recorded as 'other'. 45.4 percent were recorded as taking no medication.

2.1.13   The service provided in Newham

Of the 3371 referrals to the service, 2735 (81%) had an assessment session (this may be a telephone or face to face assessment). Reliable referral and assessment dates were available for 2678 (97.9%) of these.   The median waiting time between referral date and first contact was 6 days (mean 18.76, s.d. 30.61; range 0-360 days). For 53% referrals (n=1417) an assessment took place within one week of the referral date, 64.1% within two weeks and 84.7% within six weeks. The waiting time between referral and assessment increased over time from the inception of the service until the end of 2007 (mean 40.34, sd 38.45) after which there was a substantial decrease (mean 5.94, sd 13.22). This marked improvement was sustained to the end of the audit period, as a result of the change in service model to telephone assessment, and the appointment of an administrator to book appointments.

---

[23] Duration of current episode data were available for 57.4 percent (n=1775) of all referrals. Data were available for time since onset but were only completed for 19% of cases and therefore were not used in the analysis.
[24] PHQ9: 16.00 (5.94) vs 13.85 (6.25) t(2330) = -8.36, p<.001; GAD-7: 13.45 (5.20) vs 12.53 (5.44) t (2318) =-4.10, p<.001.
[25] PHQ9:15.12 (6.19) vs 14.08 (6.33) t (2475) =-4.03, p<.001; GAD-7: 13.20 (5.41) vs 12.29 (5.30). t(2456)=-4.04, p<.001.
[26] (14.78 (6.45) vs 13.85 (6.24) t(1987)=-2.97, p=.003)

Of the 3371 referrals 1468 had one or more intervention sessions (1371 people, 97 re-referrals). Reliable referral and intervention dates were available for 1454 (99.0%). The median waiting time between referral and first intervention session was 47 days (mean 60.61, sd 50.85); range 0-323 days.)

Of 2811 completed referrals, 1663 (59.2%) ended at or before the assessment session. The majority either did not engage with the service (574; 34.5%) or were not suitable for the service (570; 34.3%). This figure includes people 'signposted' to other services, such as the systemic service from February 2007, which by February 2008 had assessed 139 people. The local service estimates that overall 16-20% of referrals were referred onwards following initial telephone consultation.

When compared with people who attend their first intervention session there is a tendency for the young to end contact with the service at or before assessment; 61% of those in the age group 16 to 39, ended contact with the service at or before assessment compared with 57% of those over the age of 40 (X^2 =4.05, p=.04). A significantly higher proportion of those on benefits ended treatment at or before assessment (47.6% vs 40.5%, X^2 =9.48, p=.002). Those who ended treatment at or before assessment had lower anxiety as measured by the GAD (11.91 sd 7.77 vs 13.14 sd 5.21), there were no differences in levels of depression as measured by the PHQ nor were there differences by gender, ethnic group, duration of problem or waiting time (p=.12), gender or ethnic group

The average clinical time spent in contact with the patient and administering an intervention is 5 hours. When all contacts are included this rises to 7 hours 37 minutes clinical time.

In Newham, step 2 interventions included guided self help, computerized CBT, individual and group psycho-education, exercise, social support, and counselling, of which guided self-help was the most frequent intervention. Step 3 included brief CBT (max 8 hours), individual CBT (max 20 hours), CBT group and medication, of which individual or brief CBT were the most frequent. For people with one or more intervention session (n=1148), 523 (45.6%) were assigned to step 3 and 548 (47.7%) to step 2 intervention, of which 129 (23.5%; 11.2% of total) were stepped up later. The number receiving CBT was therefore 652. Those whose current episode was <1 year were more likely to receive a Step 2 intervention compared with those whose current episode was 1 year or more (χ^2 = 8.67, p=.013).

Waiting times were significantly longer for face to face compared with telephone assessments (mean 29.50 (38.33) vs 14.27 (25.51); median 19 vs 4 ; t=144, p<.001), confirming that the change in assessment policy successfully reduced waiting time. There is no difference in waiting times between ethnic groups or between self and GP referrals but those referred by Pathways to Work had a substantially lower waiting time (mean 3.48 days; median 3 days), although this is probably accounted for by these all being after January 2008, when waiting times for everyone were low.

When the service was established there were two Assistant Therapists and seven[27] Cognitive Behaviour Therapists. Over the audit period eight assistant therapists/case managers and eight CB therapists left the service, whereas eleven case managers and nine CB therapists were recruited. The staff complement at the end of the audit period was five case managers and ten CB therapists.

Employment across both sites

In addition to the downloaded data from clinical records in Newham and Doncaster, we conducted a postal survey of patients identified in general practices within Newham and Doncaster as eligible for the IAPT service. From this survey we have data on the employment status of 435 people who completed an initial questionnaire and a second one four months later. (For full details of sample recruitment and characteristics, see Chapter 6).

 In this section, we summarise information about employment status drawn from all data sources (Doncaster and Newham downloads and the Sheffield postal cohort) in order to describe the overall picture (see Appendix to Chapter 2 for data table)

• Between one third and one half of the people in the various different samples were employed (either full- or part-time) at the point of first contact (range = 45% - 51.6%).

• Around one third were unemployed and on an out of work benefit (range = 29.8 – 36.6%).

• Fewer than 1 in 20 were full-time students (range = 1.1 – 4.9%)

• Slightly more than 1 in 20 were retired (range = 3.5% - 8.1%)

---

[27] There may have been nine CB therapists, as there is a discrepancy between the staff list and the information downloaded from the IT system.

- Slightly more again were full time homemakers or carers (range = 8 – 13.6%)

- Across the demonstration sites, small increases in proportions working full-time were recorded (6.5% in the Newham service data, 3.4% in the Doncaster service data). Amongst those who were unemployed at first contact, 6.1% in Newham obtained full-time employment, and 3.6% part-time by the time of last contact.

- There was little change in part-time employment rates.

- The proportions registered unemployed reduced slightly in Doncaster (from 33.5% to 30.9%) and increased slightly in Newham (from 12.1% to 16%). The increase in Newham is in part due to an increase in the proportion of people on Incapacity Benefit who describe themselves as 'actively seeking work' at last contact compared with first contact. This suggests a shift in employment outlook for some.

- Amongst those who were in employment at both first and last contact in Newham, fewer were taking time off sick by the time of last contact, and several had changed jobs.

- 6.8% of patients in Newham moved from employment to benefits between first and last contact; 9.7% moved from benefits to employment (either full- or part-time).

We also investigated changes in benefit off-flow rates using DWP data for the period November 1999 to May 2009. Overall, there were no discernible changes directly attributable to IAPT.

Discussion

What do these data tell us about service quality in terms of accessibility, equity, coverage, acceptability and efficiency?

Although 'next day' appointments were the exception rather than the rule, and there is evidence of waiting times increasing over time, the services were, by usual NHS standards, highly accessible and have sustained this past the end of the demonstration period. Waiting times were short compared with pre-IAPT psychological services. For example, the median wait of 3.6 weeks in Doncaster compares favourably with previously recorded waiting times for psychology, CBT and counselling services in Doncaster, where waits were typically in months, sometimes many months, rather than weeks. In the second phase of Newham's service, telephone contact and triage was made at

a median of six days following referral, and the median waiting time between referral and intervention was over six weeks, which also compares favourably with pre-IAPT access.   However, this may not be a fair comparison, as the types of services being accessed are so very different and adding this degree of extra capacity into any system would be likely to reduce the waiting times.

It is hard to give a full account of service equity in the absence of data on socio-economic and educational status, but there is evidence that the service is being accessed by unemployed people and benefit claimants as was intended. Some people move into employment after using the service and there seemed to be an increase in people actively seeking work. However, the proportions are small and in relation to DWP data do not obviously alter the benefits off-flow rate for the area as a whole. Men, people over 55 and people from ethnic minorities did not access the service to the same extent as women, people under 56 and White people or as much as would be expected from Doncaster population statistics.

To what extent was the original service vision realised? In Doncaster, the aim to provide a high volume, low intensity service was successfully achieved: in May 2006 before the service began the aspiration was to provide 412 referrals per month extra capacity, and referral rates peaked at 385 per month, many more than traditional psychology services of similar size. Stepped care was implemented, with some people being assigned to step 3 from the outset and others 'stepping up' or 'stepping down' later. The basis on which the decision was made to assign to step 3 immediately is not clear, as it is not on the basis of severity. The CBT service itself was low volume. The aspiration to design a 'whole system' service with referral pathways from occupational health services, local authority staff, job centre plus staff, condition management programmes and the voluntary sector was not realised, with very few self-referrals or referrals from non-NHS services.

The original vision for the Newham service was only implemented for six months and was not fully realised or evaluated, because of the Department of Health requirement that a high-volume service be implemented in order to bring down unit costs. It was felt within the service that the long history of providing one to one therapies in Newham made the introduction of low intensity interventions difficult because patients and referrers already had experience and expectancies of face to face high intensity therapy. This imposed change in service policy led to the Newham service adopting a care pathway more similar to the Doncaster model, but without the same organisational commitment and staff enthusiasm.

The stepped care model that was finally implemented was successful in achieving the aspiration of a balanced hybrid between low and high intensity

therapy, with a high proportion of self-referral.   This enabled Newham to offer a higher volume CBT service than Doncaster. A major challenge in Newham was to make the service genuinely accessible to people from Black, Asian and other ethnic groups including White ethnic minorities (e.g. people from East European or Kurdish backgrounds). Although people from these groups, particularly from the Asian ethnic group, were slightly under-represented, huge effort was expended in reaching out to these communities, and to retain people in the service. In addition the referral pathways created by the service, in terms of self-referral and 'pathways to work' referral opened access to ethnic minorities. There is no evidence of major inequity of access and the ethnic mix in the Newham service represents a considerable achievement.

3 Organisational aspects of the IAPT demonstration sites

Background and introduction

The IAPT demonstration projects were designed to test the feasibility of setting up a new organisational model of providing psychological therapies, and to demonstrate what could be achieved by doing so, in terms of improving access to effective psychological therapies. The organisational processes by which these services were established, developed and sustained are therefore important to understand, to learn which factors facilitated or hindered successful implementation of this innovation.

The two most striking features of the task which faced the organisations were the complexity of the innovation and the short timescale for implementation. In the context of mental health service delivery systems, IAPT uniquely combines several factors.   The original model comprised:

- a new service format;

- training and recruiting new staff groups;

- a service which is highly flexible and provides rapid access;

- an emphasis on non-traditional approaches reaching out to non-traditional users of psychological therapies;

- delivering therapies in new ways (e.g. by telephone);

- extensive partnership working across both NHS and non-NHS organisations;

- liaison with a wide range of stakeholders and gatekeepers;

- parallel implementation of new IT systems; and

- an unprecedented level of repeated measurement of patients' progress and outcomes.

The organisational challenge of implementing a programme of this complexity was increased by the very demanding timescale within which results were needed. The interval between first commissioning these services and their

being 'open for business' was seven months, (or even less than this if confirmation of funding for the innovation is taken as the start date). It was only a further six months between start-up and the Department of Health's deadline for outcome data from completed cases which could support the business case to H.M. Treasury for IAPT roll-out.

These challenges required a very large number of practical obstacles to be overcome in terms of staff management, training, recruitment, IT system implementation, accommodation, referral protocols, engaging stakeholders, etc. In addition to overcoming problems with these practicalities, it would be expected that the IAPT innovation inevitably creates a challenge to the existing organisational culture within teams and professional groups, which could react with scepticism, resistance or lack of co-operation. A new service delivery method could be seen to threaten or undermine other groups' work in this field, so that grafting the innovation onto existing service systems requires leadership and negotiation skills. We found a previous SDO report by Greenhalgh et al. [10] useful in reviewing evidence for the successful introduction and sustainability of innovations in the NHS. This found seven key areas which are influential, and these informed our thinking at the outset of this organisational case study. These are:

- the attributes of the innovation itself;

- the adoption process as engaged in (or not) by individuals;

- communication and influence (including the impact of opinion leaders, champions, boundary spanners and designated change agents);

- the inner (organisational) context (including structural determinants of innovativeness, receptive context for change in general, absorptive capacity for new knowledge, and tension for a particular change);

- the outer (extra-organisational) context (including inter-organisational collaboration and networking, prevailing environmental pressures such as external competition, particular policymaking contexts and streams, and proactive linkage initiatives);

- the nature of any active dissemination campaign (which incorporates the general principles of social marketing and knowledge construction); and

- the nature of any active implementation process (which incorporates the general principles of effective management in a changing environment).

Aim and objectives

The aim of the case study work at each site is to provide an understanding of how the service was implemented and any learning relevant to the roll out of IAPT across the NHS as a whole.

Objectives:

To provide an organisational analysis of the process of implementing and maintaining the required changes in service delivery at the two demonstration sites.

To evaluate the extent to which the organisations were able to fulfil their stated objectives, highlighting helpful and hindering organisational processes and issues.

Method

The design of the case studies drew on the methodological principles for examining organisational change put forward in Pettigrew [11,12]. This advocates the adoption of a pluralist approach which allows for different perspectives and contexts, studied over the duration of the change/implementation process. This section summarises the main methodological steps.

3.1.1 Scoping exercise

The initial phase in each case study was to undertake a detailed scoping exercise to identify key stakeholders and players at different levels within each demonstration site creating a pool of stakeholders from whom interviewees could be drawn and ensuring multiple perspectives were represented.

3.1.2 Sample

In line with the proposed change management framework described by Pawson [13], each case study consisted of work at a number of levels.

A total of 57 people were interviewed between November 2007 and September 2008 at each of the demonstration sites. A further 10 participants (nine interviews) were interviewed during the latter stages of the demonstration period and as both sites moved to local funding arrangements (see Table 9 for details). These latter stages of the case studies included

interviews with participants from the comparator sites as they prepared to implement IAPT.

Table 9. Interviewees by role

	N Round 1	N Round 2
Strategic managers (includes PCT & MH Trusts & Chief Execs, partner orgs)	17	5
Operational managers (IAPT & partners)	10	4
Other stakeholders (GPs, proj consultants)	8	1
Front line staff (therapists & case managers)	18	
National IAPT leaders	4	
Total	57	10*

*Includes one joint interview giving nine separate interviews

3.1.3 Qualitative interviews

The interviews were partially structured by a topic guide appropriate for the type of participant (Appendix to Chapter 3) which provided a range of issues for exploration whilst leaving the informant free to talk about their experiences and perspectives without intrusive questioning. Interview duration varied between 45-90 minutes. Interviews were digitally recorded and transcribed.

3.1.4 Data analytic method

The data analytic approach was derived from the Framework method, in that interview elements were first indexed, coded then analysed thematically, with recurrent themes identified from the transcripts and with reference to the research questions identified in the proposal. A team of five researchers undertook the analysis (GP, GH, JR, EC, JC, RH) with support from one other (KD).

Initial indexing was undertaken on eight interview transcripts, which yielded 32 themes, derived consensually after two or more researchers independently

identified themes from each interview. These 32 themes were then examined by the full team and grouped into seven higher-order themes. The individual themes within these seven groupings were used as the coding framework. Appendix to Chapter 3 provides details of the individual themes identified. Remaining interviews were then coded by a single researcher.

For the second round of interviews the existing coding frame was applied to allow comparison of themes. The original 32 themes were found to be too specific for data from interviews where participants were reflecting on the progress of IAPT over the previous three years, or, as was the case for interviewees from the comparator sites, where IAPT had not been in place. However, the higher level codes were found to be sufficiently comprehensive as to allow the coding of the interviews at that level.

As for the first round of data collection, each interview was analysed by a single researcher and the coding units collated for each top level theme separately within an Excel spreadsheet. The key points for each top level theme were then extracted and written up by members of the team.

Findings

The higher order themes derived from the analysis of the organisational case study data were:

- Top-down drivers

- The innovation

- Stakeholders

- Organisational systems

- Implementation

- Job Characteristics

- Sustainability

The overall findings are summarised here by higher order theme (further detail of the earlier case study work is available at www.shef.ac.uk/cpsr/projects/iapt.html.

Finally two overarching sets of findings which emerged from the data are described:

- Partnership working

- Sustainability and local development

3.1.5 Top down drivers

This theme included five sub themes all relating to external pressures and the wider context, particularly the political context, within which the demonstration sites were set up.

External pressure to deliver and lack of control

- External context (shifting goal posts, changing demands, local & national tensions)
- Tactical, political nature of the demonstration projects
- External relationships with VIPs
- Impact of routine data collection on staff and service.

There were a number of external forces driving forward and shaping the IAPT demonstration sites. Top down drivers reflects both the nature of the political drivers and the practical implications for both sites.

At the early stages of the case studies many comments were made about the tactical/political nature of the demonstration sites and the consequent implications in terms of pressure to deliver and fear of the consequences of failure. The fact that this was a politically high profile innovation and that many hands were felt to be guiding the emerging policy in this area, meant for the demonstration sites the sense of constantly shifting goalposts, changing demands and success criteria and an inability to gain clarity about fundamental aspects of the demonstration sites, such as confirmation of funding, at the same time as the sites themselves needed to sign contracts for provision of services. This experience was similar at both sites, although for Newham there was the added pressure of needing to re-design the service after the first six months of provision.

> *"There was an 8% cut, then a 20% cut, then no cut at all, so we had to go from planning a drastic cut in the service through to un-planning the cut and proceeding."*

> *"The national emerging IAPT programme kept changing its emphasis slightly in terms of…employment was always a big focus, but then they changed the goalposts and it wasn't much really."*

By the time of the second phase of data collection there was an overall sense of relief expressed by both sites to be free from the 'straightjacket' of the demonstration period. Whereas previously the main service drivers were felt to be the DH, the emphasis had now shifted to local commissioning practice. Whilst on the one hand this was felt to provide an opportunity to take stock of what was needed from a commissioning perspective, the split between commissioning and service provision and the advent of Practice Based Commissioning were felt to exert strong and sometimes not fully aligned influences over the direction of the IAPT service. Additionally, it was evident that although there may be more freedom at a local level in the structure of IAPT services, there was little choice regarding its overall continuation.

> *"It very quickly emerged that, regardless of any commissioning intentions we might have had, we were told we had to continue to fund the service"*

The comparator sites appeared to be experiencing many of the frustrations voiced in the early stages at Doncaster and Newham, reporting lack of clarity around performance targets and funding, as well as expressing concerns about the constraints of the IAPT service model.

3.1.6 The innovation

The innovation theme captured comments about the nature of the IAPT model and the way it operated. Comments fell in to five main areas:

- the design of the service including stepped care, case management, referral pathways, exclusion/inclusion criteria and the aspirations to flexible and community based approaches;

- systems thinking covering holistic vs traditional 'silo' thinking and how well patient journeys/care pathways work;

- the physical location of the service;

- new ways of working; and

- engaging gatekeepers/referrers including issues around working with the range of possible service collaborators including GPs, businesses, occupational health services, Jobcentre Plus, community and voluntary sector organisations.

During the early phase of data collection the aspirations of IAPT were clear in addressing mental health inequalities, breaking down barriers and creating a service that supported individuals in remaining active in society and in their

community. Moving away from traditional clinical delivery methods was cited as a key way of avoiding some of the stigma attached to mental health issues:

> *"We were determined that we did not want this to be seen as a health or illness service, we wanted it to be very much embedded with the business community".*

And in overcoming shortfalls in support for those in employment who are finding it difficult to cope and in providing access to those in previously hard to reach communities i.e. those from black and ethnic minority communities and non English speakers:

> *"…what can we do to break down those barriers and create incentives that help people to remain active in society and in their communities, which to a large part does include employment, but not exclusively, and still provide them access to health care?"*

Much of the early data gathered on this theme concerned partnerships and the centrality of relationships in getting the new services up and running. Many participants commented that in hindsight more time and a developmental approach may have been more sensible.

Participants expressed the wish for the new services to move from 'silo' thinking to being a patient focused and whole systems approach, so that patients should experience continuity of care. To achieve this there was recognition that there needed to be a more flexible approach and one where problems with patient care were monitored, recognised and actioned.

Accommodation was seen to facilitate and embody the culture of the new service. It was a powerful symbol of what was new to both the public and other government services. Analyses highlighted issues around accommodation that ranged from practical concerns (such as lack of space) to the vision for the service and perceived public image. Some of the arguments about location centred on NHS versus community sites, and primary versus secondary care integration. As might be expected, actually finding adequate accommodation was a central and ongoing concern at both sites.

Another issue relating to accommodation, but highlighted more broadly in interviews, concerned computer access and IT systems. The use of a variety of locations meant that access to computers was often problematic, causing great difficulties for front line staff providing a service heavily reliant on IT. There were also confidentiality issues linked to this; for example case management supervision using the IT system having to take place in

cramped open plan offices, where conversations could be easily overheard (even by patients at the other end of the telephone lines).

The speed in which the services were implemented meant that preparatory discussion and agreement about important aspects of the innovation such as referral pathways was not possible, which led to opposition from other professional groups.

At the stage two data collection, most comments from the demonstration sites related to improvements and expansion of existing services and many of the earlier problems had been resolved. Doncaster had moved to much better offices; whereas suitable and sufficient accommodation remains an issue for Newham.

Difficulties between primary and secondary services remained an issue at both sites and at Doncaster approaches such as joint clinics were being explored as a way to heal the rift between the services. At Newham the issue of how secondary services would work with IAPT remained prominent and there was acknowledgement that the service needed to be streamlined, that the IT system needed improvement and the practical working of a single point of entry into mental health services needed work.

For comparator sites, who had yet to encounter many of the practical challenges experienced by the demonstration sites, there was a general feeling that the national guidance was helpful, but too prescriptive. There was a strong sense of the need to adapt IAPT to the local context.

3.1.7 Stakeholders

The stakeholders code represented a range of views on the new IAPT service from, at the broadest level, societal perspectives including tackling stigma around mental health issues, public image and awareness of the service, to more specific concerns regarding for example, the acceptability of new ways of working. This theme reflected some of the complexities of inter-organisational partnerships, including who is 'in' or 'outside' the partnership, dynamics between primary and secondary care, ownership of IAPT, the quality of relationships and the extent of collaboration. This was also an area where comment was made on service user and third sector involvement. The theme consisted of five sub themes:

- Inter-organisational partnerships
- Service user and representative (3rd sector) organisation perspectives

- Societal perspectives

- Use of external expertise to guide, train and mentor

- Tension between public and private sector ways of working.

Comments on stakeholder involvement in the partnerships tended to focus more specifically on the experiences of and necessary conditions for effective partnership working. At both sites some examples were given of this working well at strategic and operational levels. However, there were many more examples of dissatisfaction with the way the partnerships were working and recognition that there is a considerable difference between partners signing up to working collaboratively at the strategic level (eg developing the shared vision) and making it work operationally on a day to day basis.

A particular example of this concerned referral pathways. Getting local GPs to be positive about the new service was seen as paramount, but often problematic. In both sites there were some GPs who valued and used the service and some who were more reluctant or never did. Having a GP champion for the new service was seen as vital and using GP forums for engaging practices was found to be useful.

Organisational perspectives on service user involvement during the early stages of the research were mixed. There was also recognition that the level of patient involvement in the design of services was modest – users were not centrally involved in the design of the service. It had been difficult to maintain contact, however there were signs that the situation was improving.

> *"there's been a significant change there because X left …the patient forum group struggled for a bit. We have a core membership of three now and we have a new patient advisor and they're wanting to change the name and everything, put their own stamp on it so we're hoping – that's just sort of been picked up again*

Speed and ease of access were seen by the service user representative as the great advantage of the IAPT model over previous service configurations, but this was tempered by more negative views of telephone rather than face to face contact, the frequency of measures and some components of low intensity work such as bibliotherapy and CCBT. Additionally, some specific aspects of service design, such as initial assessment by a high intensity worker, with follow up by a case manager for guided self help were found to be problematic.

Overall, interviewees reported high levels of patient satisfaction with the service, however there was recognition that measures of satisfaction could be

biased as feedback is given directly to the case manager. This means a bias in favour of those who complete treatment (as opposed to those who drop out) and a situational demand as clients provide feedback directly to the person providing the service, making it more likely that they will emphasise positives and play down negatives.

By the second round of data collection, much comment in relation to the stakeholders theme again focused on the complexities of partnership working. Primarily this concerned the difficulties, following changes in commissioning practices, of organisations working collaboratively over IAPT whilst at the same time competing directly for funding and contracts in other areas. There also appeared to be little resolution of earlier difficulties with partnership working - the legacy of the implementation time-frame. As a result, stakeholder collaboration in delivery of IAPT remained problematic at both sites. This is discussed further in section 3.5.1.

It was equally clear that the involvement of service users, peripheral in the early stages of the demonstration period, was difficult to sustain, with too much reliance on one particular individual.

3.1.8 Organisational systems

Data in this area related to the structures and systems in place in the organisations responsible for the direct delivery of the IAPT service. Themes included:

- contractual, financial and organisational system constraints on service delivery;

- resource issues;

- workforce issues at the organisational level; and

- issues around system capacity.

Initial funding delays contributed to difficulties in setting up services at both sites. Delays in setting a budget and political uncertainties made it difficult for both sites to make appointments, with the consequence that when funding was confirmed there was intense pressure to appoint to vacancies to meet project deadlines:

> *"Having the funding confirmed was very much a stop-start activity which doesn't sit at all with setting things up so that was a bit messy."*

The project status of the demonstration sites appeared to cause particular problems during the early stage of data collection, with both sites reporting little clarity as to governance and accountability arrangements. There was a feeling, expressed at both sites, that the nature of the demonstration sites (in particular their high political profile) led to more complex arrangements with regard to decision making and frustrations amongst partner organisations who felt they could have managed the arrangements better.

Both IT and admin systems caused difficulties for the delivery of the service. Both functions were felt to face complications due to the split site nature of IAPT delivery, for example IT access was not possible at all locations where sessions took place, meaning the need to keep handwritten notes to transfer back on to the system at the centre. There were more fundamental challenges at Newham where, for example, discussion over the intellectual property rights to the newly developed IT system risked delaying implementation. As it was IAPT staff had to keep paper records for several months and then much therapist time was spent entering all the data on to the new system.

The philosophy behind the recruitment of case managers was to include people without traditional professional backgrounds and people with personal experience of mental health difficulties, or of caring for someone with mental health difficulties. Being a demonstration site and deploying a 'new' type of employee meant that there was not a clear precedent to follow in terms of agreeing appropriate pay banding for different levels of therapist.

Broader contractual issues aside, no amount of resource could overcome the difficulties encountered in trying to recruit and retain experienced CBT therapists. This was particularly problematic at Doncaster where historically low levels of CBT provision were felt to compound the problem.

Recruitment and retention difficulties were created due to the ongoing funding issues. Recruitment of therapists started and stopped at both sites at different times, dictated by the degree of funding uncertainty. This meant that recruitment didn't keep pace with turnover or demand. Additionally uncertainties for staff up to March 2008 (a month before the end of the demonstration period) led directly to the loss of experienced employees to newly set up IAPT services who could offer better terms and conditions and a greater degree of job security. This pattern was still evident at the second round of data collection where for example Newham was struggling to fill low intensity roles as competitor IAPT services were recruiting their low intensity workers into high intensity roles in the new services.

By the second round of data collection financial uncertainties continued to cause problems at Newham. This was in part because the second wave of IAPT had been commissioned to run across the Borough following the end of the demonstration period. For service managers the funding position was still not clear at follow up however, and there was concern that large numbers of new staff were being appointed without any clear apportionment of the funding between the PCT and the Specialist Trust, to employ low intensity and high intensity workers respectively.

Additionally the administrative support team at Newham were under review, and a number of vacancies there couldn't be filled until the outcome of the review was known. Current staffing levels were viewed as untenable, and too thinly stretched across the gamut of psychological therapy services.

At both sites the services suffered from the loss of experienced staff and their replacement with newly qualified or trainee personnel, with knock on effects for waiting times, service capacity and workloads.

> *"We lost 50% of our low intensity case managers at some point and so there was a huge waiting list built up on low intensity and we had to get a lot of people – we stepped them up to high intensity".*

Issues of wait list management were being tackled with a range of strategies at both sites. At Newham people waiting for high intensity therapy were invited to attend one off 'workshops' to get them involved in doing something, and some therapists were also offering group work as a way of keeping people engaged whilst they waited for one to one therapy. There was a fear that without this, the DNA rates would start to get too high. There were also plans to tighten up on the existing system of case reviews to ensure that therapists and supervisors were able to promptly identify people who should be stepped up or discharged, to create space for taking new people on.

Some of the discrepancies in waiting times for high intensity CBT which existed previously in Doncaster between secondary care and IAPT appeared to have been tackled at the time of follow up. Secondary services had invested heavily in reducing their waiting times to match IAPT's – although it was not clear how sustainable this work would be, given continuing shortages in qualified CBT staff, and funding. Group work led by high intensity therapists was also being deployed at Doncaster for those on the waiting list.

At Doncaster, concern about inappropriate referrals being 'bounced back and forth between IAPT and secondary services was still an issue at the time of

second interview – a concern that would be in part addressed through ongoing work towards a single point of access to the service..

3.1.9 Implementation

The implementation theme captured comments made by participants directly about the implementation process itself. Analysis revealed the following areas of comment or concern:

- Implementation process and change management;
- Vision and leadership;
- Grafting an innovation on to an existing service;
- Winning hearts and minds;
- Planned and intended consequences; and
- Unintended consequences and knock on effects outside IAPT

As might be anticipated given the challenges already identified arising from the implementation timescale, there were many general criticisms of the way implementation was handled from lack of co-ordination amongst the many partners involved to the speed with which major decisions had to be made without proper opportunity for reflection or planning. The issues were captured eloquently by the following participant:

> *"Ok, I think there's one really clear lesson - and it is a bit of a mantra of mine - I think we are very quick in the NHS and we did this with IAPT, to come up with an idea, and jump to the solution, decide what the answer is, and we specify what we think the answer is, and then we start implementing, and then during implementation, we find all the issues and we take a long time to get to the point where it's actually operating in the way we want. And I think what we really needed to do with IAPT, and I think we're trying to do this time, and with lots of other things in the NHS, is we need to take much more of a developmental approach to getting to the answer…Yes. And not see that as wasted time – you know like you've been working on this for six months, have you not got a specification yet? And I keep saying show me where, anywhere in the commercial market, somebody takes a new product to market and develops it in less than six months? And the NHS forces itself into, we force ourselves because of political pressure, into jumping to the solution, without doing the testing phase…So why do we not get the speed of uptake for us to be able to demonstrate the viability of the model? Because we spent literally no time communicating with GPs about the*

change, before we changed it! Why do we have the issues about the relationship with CMHTs and the movement of clients between the different parts of the system – because we have never sat down and simulated what that would be like. So, I do think that IAPT was an example of that, and that was because of the national push".

Vision and leadership were also felt to be compromised. Participants reporting that IAPT produced tensions regarding the service model and lack of understanding, communication, clarity or guidance about how it would work. This included confusion about key aspects of the service such as who it should be targeting. The situation in Doncaster was somewhat mitigated by the considerable body of work that had been undertaken around planning services for mental health prior to the advent of the IAPT pilot, however this did not completely counteract the difficulties encountered.

The importance of specific individuals who had taken on leadership roles and driven the process was commented on at both sites, as was the recognition that this can pose particular problems should that individual leave.

Many of the day to day implementation concerns revolved around the practicalities of grafting a new innovation on to existing services. Invariably existing services felt threatened by what was seen as 'a brand, spanking, new, all singing, all dancing…service' which had appeared out of nowhere, with little or no preparatory ground work. Between both the old and the new services there was a sense, reported by several participants, of a lack of understanding and a lack of willingness to understand. The language in these parts of the interviews is telling: Tribalism; territorialism; protectiveness; a few individuals with grievances; rivalry; suspicion; caution; threat; hostility to new ways of working; scepticism… The message from this being that challenges experienced at the strategic level were impacting no less painfully on the frontline.

Some participants did discuss facilitators to the process. They noted the need for ongoing dialogue, negotiation and discussion, noting the considerable management resource implications.

Closely aligned to introduction of the new service, the 'winning hearts and minds' theme captured discussion of the work done to 'sell' IAPT both internally and externally. As a concept it is clearly intertwined with engaging gatekeepers and inter-organisational partnerships. Participants identified a range of activities undertaken in the promotion of IAPT including: Marketing; education sessions; awareness raising sessions; liaising and building links with potential internal and external partners; informal networking; more

formal partnerships; and collaborative working. A variety of methods and approaches were used from presentations and meetings to running conferences and stakeholder days. It is clear that a considerable amount of work and resource was expended on promoting the IAPT service and overcoming barriers. There was also recognition at the early stage of data collection that further in depth work was required. Despite the obvious implementation difficulties encountered there were reports of the many and intended outcomes that, even by the early stage of data collection, it was felt had been achieved. In particular the fact that the principles of the new model of working (eg case management, brief intervention, a stepped approach and the use of telephones for rapid access) were demonstrated to be working in practice. The speed of access permitted by IAPT was repeated by many participants and, with the wider choice offered to patients and referrers represented a significant achievement for the services.

Additionally the use of the information systems and sessional measures meant that workers themselves could see the changes in patients more easily.

Overall there was a sense of success as demonstration sites in that:

> *"I think demonstration, I am quite happy with the phrase 'demonstration site' because I think that's what we were able to do … demonstrate to other parts of the service and to the Department that a model of Improving Access to Psychological Therapies could be implemented and could work".*

Inevitably there were areas where intended consequences had not been achieved. At both sites the model was 'not quite right' and there were ongoing adjustments. Primarily the lack of joint working and failure in systems between old and new services was disappointing and fell below what had been hoped for.

At Doncaster links with employment services were still an area to be developed although a more integrated approach had been achieved at Newham which seemed to be working well despite initial teething problems.

Whilst both sites reported positive unanticipated consequences of the introduction of IAPT (examples included: GPs screening for depression/anxiety as they realise they are able to manage moderate depression in the community with use of IAPT; and Case Managers based in Sure Start areas doing beneficial outreach and other work eg. postnatal depression sessions, engaging families who previously hadn't engaged with psychological therapies) there were also unintended negative consequences,

particularly the difficulties created between staff in new and existing services and the lack of clarity identified around referral pathways in general.

One of the key lessons from the demonstration period identified by Doncaster was the need for sufficient management resource to run the service. At follow up, the addition of two deputy manager roles proved a helpful development in supporting the main service manager and enabling the smoother running of the service.

By the second period of data collection there had been a striking shift in the types of comments captured by this code, with a move away from comments focussed on intra and inter service issues to a much broader reflection on the service as a whole. The focus had moved much more on to patient needs and balancing the service accordingly.

There were changes to the original service specification. As IAPT matured, other projects and services were 'bolted on' and the scope of the service changed and expanded. This is an ongoing process and changes continue to be made in staff roles and structures to achieve a more integrated response to the needs of more complex cases or people in crisis.

3.1.10   Job characteristics

The job characteristics code focuses on the direct experiences of people working at the IAPT demonstration sites. At the individual level the theme captures experience in the job including job satisfaction, perceptions of job demands, coping mechanisms, staff turnover, job design & problems with job role. At the team level, the theme reflects data on intra-IAPT collaboration or conflict, support and organisational culture/climate. It includes sub themes of team working, team support, peer support, supervision support and collegial relationships. Finally this code includes the sub theme of management style and competence including line management, project management and delegation skills.:

- Job experience
- Intra-IAPT collaboration
- Job design
- Management style and competence

During the first round of data collection, the considerable job satisfaction experienced by case managers and therapists was evident despite often trying local conditions:

> *"I do believe that we've made a difference to a lot of folks' lives which is a good reason to get out of bed in the morning"*

However, the operational difficulties at the early stages of the demonstration sites as already discussed translated into challenging situations for many interviewees:

> *"I have witnessed people speaking to each other in a way that I have never witnessed people speaking before."*

> *"My … colleagues and the manager will come back almost in tears – or they will go there with cases…and find that there are four people there from XXXX lined up ready to attack them. Really, it was that unpleasant."*

In addition to inter and intra service conflict for employees providing the new IAPT service there were a number of demanding aspects to the job. High caseloads and the implications for time management were frequently referred to For many case managers, in order to meet the stipulated 20 clinical hours per week they had to book appointments back to back. Booking clients in this way meant that case managers were dependent on drop out to fulfil other core parts of their job. This was generally manageable until:

- everybody booked answered the phone/turned up for appointments, in which case there was no time to fit in other important tasks that need doing or to have a break

- there was a difficult case which required attention, the most commonly cited problem was a client disclosing suicide risk. This had a knock on effect for all following appointments and there was often nobody available to ring round and postpone the next appointments

- A colleague was off sick, or somebody else's telephone calls needed covering because of falling behind due to a crisis

Support from peers, line and clinical management was paramount in helping frontline staff to cope with the demands of the job. In general, participants' comments from both demonstration sites indicate that supervision is working well and is highly valued:

Project 08/1610/154

"It is lovely to have positive role models around you and see people that you respect when you see them in action … my clinical supervisor is amazing, really experienced, she's awesome, really inspiring."

During the early phase of data collection it was clear that the desired level of, or access to supervision had not always been available. Reasons for lack of appropriate supervision varied, including increased demand for supervisory support (e.g. through 'inappropriate' referrals) resistance to the role from senior staff, under-resourcing of the supervisory role, issues with the peripatetic nature of the service and lack of clarity around who to contact, all of which can have a negative impact on staff:

"It's a bit stressful not knowing who your line manager is and who you are supposed to be calling. I mean I know XXX is happy to do it but I think that we worry about the fact that a lot of us call him, put a lot of pressure on him, and if we are supposed to be calling somebody else then that's not fair on him is it? We're not really clear on who we are supposed to be calling."

At both demonstration sites some formal measures had been taken to address the issue, either through restructuring the team to create duty manager rotas or increasing (through appointment) the resource available to fill this role. However, at both sites participants gave examples of phoning colleagues at home out of hours and ad hoc arrangements including a senior colleague extending their role informally.

Project management is the key to the way that front line staff experience their roles. There was a lot that was new about IAPT and therefore new for the project management teams involved in the delivery.

The management resource required by the demonstration sites was initially underestimated. Failure to bring in an experienced project manager at the outset, and the speed with which the services had to be set up, led to basic and/or crucial operational management tasks being neglected. Frequently mentioned issues included lack of budget control, IT system contract problems, pay problems, inadequate workspace, breakdown in links between HR, IT and Finance and fundamentally, lack of clarity about how to resolve day to day operating issues around workloads, supervision and referral issues. Both demonstration sites were to resolve these issues by bringing in experienced project managers at a later stage in the project. At both sites the late arrival of adequately qualified/experienced project managers resulted in the role initially being one of troubleshooting.

By the second stage of data collection both sites were experiencing uncertainty as the pilot period ended and tendering processes for the services were considered. In the end this did not happen in Doncaster. The interviewees also thought that the structures and systems in Doncaster were now established and the service no longer depended on particular individuals. There was however some frustration at the lack of career structure and development at Doncaster.

The Newham service, at the time of data collection was just clarifying the new contract for services after a protracted commissioning process. The service is still likely to expand, although a lack of information at the time made planning difficult. Additionally there was frustration at the lack of clarity around managerial roles, a sense that there remained a lack of sufficient managerial resource, and with regard to the commissioning process, that although the tender was a joint bid with the PCT there remained a lack of clarity about how exactly services would be delivered and how the PCT and IAPT services would work together.

The Newham comparator had set up an IAPT service in the previous year. This new service had a clear structure with 4 teams each headed by an experienced psychologist. An emphasis on team working, good leadership and clear communication meant that much had been achieved, although some staff had moved jobs as a consequence of the new service structure.

3.1.11 Sustainability

The Sustainability theme groups all comments on the maintenance and development of the service which fall into three broad categories:

- The sustainability of the new model vs. revision to the status quo
- Plans for the future and lessons learned; and
- Commissioner perspectives

One set of comments on the sustainability of IAPT during the early round of data collection concerned the ability of services to maintain all that was new and innovative about the IAPT service and avoid drifting back to conventional ways of working. There was concern that without strong drivers in place, and in spite of genuine commitment to partnership working, that ultimately there was a risk that partners would revert to 'what they know'. Inevitably this theme is closely linked to the broader issue of partnership working which is discussed in the next section.

With regard to lessons learned and plans for the future, a clear need was expressed by different participants to examine the interface between primary and secondary care, how IAPT links in with both of these and how to "mainstream" IAPT. It was suggested that as part of this process, consideration needed to be given to a single assessment process for all services, including independent sector services.

It was also felt necessary to ensure IAPT developed in the future with regard to other systems, including Practice Based Commissioning (PBC) and the development of Foundation Trust status. More broadly, the need was expressed for future IAPT services to fit in with other regional initiatives to ensure consistency, for example with developments resulting from the Darzi review. All of this, it was felt, would require a more developmental and flexible approach in the delivery of IAPT than had been possible at set up.

Finally, in terms of lessons learned, the essential role of relationships was again identified in order to work successfully in partnership as IAPT required.

From the commissioner perspective, there was recognition of the need to learn from the demonstration sites and that in the future there would need to be development of certain aspects of IAPT. Areas within IAPT that were specifically mentioned during the first phase of data collection included stepping clients up and down, the interface between primary and secondary care, achieving manageable referral levels, better understanding of the volume of unmet need, and the possibilities of a single point of access/triage type model for all mental health services

The issue of IAPT being re-tendered was discussed and two main implications of this were highlighted, including "mainstreaming" the cost of IAPT and the implications of the purchaser/provider split in putting together a tender.

For both demonstration sites the move from demonstration to full service status brought about a period of uncertainty which at times threatened the stability of the services, particularly with rumours circulating and, as the end of individual contracts approached it became increasingly difficult to keep teams stable and motivated. In addition to uncertainty about future funding, competition for staff from newly established IAPT services was seen as a threat to sustainability.

At follow up there was a striking contrast between the demonstration sites in their experience of pressure from the centre in terms of dictating how IAPT

was operationalised as the programme moved in to the national implementation phase. At one site (Doncaster) there was a sense of being out of the spotlight and the intense pressure of being a national demonstration site dropping away. There was recognition of what the model had achieved and attention was now on sustaining and developing the IAPT services:

However, at the other demonstration site (Newham) it was felt that with pressure to continue to find funding for IAPT, they were still grappling with issues of service duplication and the failure of inter-NHS collaboration which ran counter to the locally assessed need. Consequently considerable work was undertaken to create a sustainable model for future psychological services. That said, within the context of continued funding for IAPT, there appeared to be more scope than previously experienced at Newham for adapting and developing the IAPT model.

As a result, at both sites, changes were underway to develop and extend the model or to align it with developing mainstream services. These areas are discussed more fully in section 3.5.2.

In contrast to the demonstration sites, at the comparator sites that were in the process of implementing IAPT there was considerable concern about the pressure to conform strictly to the IAPT model and whether this met locally assessed need.

Evidence on key themes

This section describes two key aspects of IAPT which emerged during the analysis of the organisational case study data.

- Partnership working was frequently commented on by participants at all levels in the organisational case studies. The type of data ranged from direct comments about the nature of partnership working and the nature of the partnership at each demonstration site to examples of how partnership working impacted on the day to day running of the IAPT services. As such, partnership working was felt to be an overarching theme which cut across the seven higher order themes described above.

- Sustainability and local development emerged as a dominant theme during the second round of data collection. This is perhaps not surprising given the combination of re-tendering of services and the

relaxation of controls around the original IAPT model at the end of the demonstration period. It is also likely to be a reflection of the issues occupying the sample population during the second round of data collection who were all in senior/managerial positions.

3.1.12   Partnership working

IAPT is an innovative programme aiming to break new ground in a number of ways: Linking NHS with employment services and employers; developing the role of case managers using guided self-help (i.e. low intensity interventions); and developing new ways of working across primary and secondary care. Delivering IAPT therefore requires collaboration between different organisations, both within the NHS and between the NHS and other sectors; for example in relation to the 'returning to work' ethos, Jobcentres and local employers.

Such partnership working is complex and requires time and resource to become fully established. At Doncaster there was already a consortium of NHS and non-NHS organisations, pre-dating the advent of IAPT, which had undertaken work to assess mental health needs within the region. At Newham, a new partnership formed specifically for the delivery of IAPT services during the demonstration period.

Despite their different starting points, partnership working proved problematic at both demonstration sites for a mixture of reasons. At the broadest level, ambiguity and uncertainty as to the constitution of the partnership, who was 'in' and who was 'out' added to the complexity of joint working. For example, the partnership was alternately seen as an intra-NHS one (between primary and secondary care) or an extra-NHS one (between NHS, business and third sector partners) and the accompanying issues were prioritised and tackled differently in each case.

Examples were given of partnerships working well at the strategic level, across NHS and non-NHS organisations, with genuine commitment to involving and using partners fully. There were also many examples given of how, despite strategic 'shared vision' there were barriers and difficulties in translating this into operational practice. The various challenges to maintaining partnership working included conflicting priorities between partners, issues on data sharing and perceived changes in the programme's priorities. There were also perceived examples of organisations wanting to show that they are involved with partnership working to 'tick the box', to

present an impression of consultation or as an expedient route to winning resources.

Service user involvement is a fundamental principle within IAPT and was seen as a key resource to aid the process of partnership working. In practice however, service user consultation was very limited at both sites, often reliant on one or two individuals and never fully realised in the way originally envisaged.

Another important aspect of IAPT was the cross partner work in relation to employment and this was an area that suffered significant challenges at the operational level at both demonstration sites during the early stages, although after initial teething problems was felt to be working well at Newham.

The changing contractual climate was also cited by participants as a challenge for the sustainability of partnership working, both within and beyond the NHS.

Such problems did not diminish when the relationships between NHS organisations were considered. At both the strategic and operational levels, there were tensions between the partners' agendas, sometimes conflict, and inevitably powerful barriers to working effectively together. Such conflict between NHS organisations at a strategic level about 'ownership' of IAPT, not only undermined the partnership approach, but impacted on a day to day operational level for staff.

Participants reported these tensions between primary and secondary mental health care providers; between therapists of different professional backgrounds or between existing services and IAPT.

On a practical level, there were not always the systems in place to resolve day to day issues affecting the running of the service, for example achieving clarity and embedding practice around referral routes and processes between partner organisations. At Doncaster, time had been spent by the clinical lead to agree a referral matrix (see Appendix to Chapter 3), clarifying areas of responsibility for the IAPT and existing services which went some way to helping different NHS providers understand how they should be working together.

Tensions sometimes escalated into a breakdown in respect, trust and professionalism between people who were expected to work together. Interpersonal hostility, particularly when combined with failure to resolve operational issues made for conditions where the use of inter- or intra-

organisational networks was difficult to sustain. Such difficulties all have the potential to limit the effectiveness of the IAPT programme to the detriment of service users.

Steps had been undertaken at both sites to address some of these issues and interviewees identified a number of factors or activities that were found helpful in overcoming distrust, misunderstandings and practical barriers between partners and facilitating genuine partnership working.

- Building on existing local strengths, either existing strategic partnerships or individuals with experience and good track record for inter-organisational liaison;

- An inclusive and transparent approach in developing collaborations.

- Clear vision and strong leadership from the PCT commissioners about the purpose of IAPT and an emphasis on collaboration between NHS, employers, JobCentres and voluntary sector;

- Clear managerial commitment to designing services around the needs of the patient or service user (i.e. user-centred service design);

- Externally facilitated meetings/development activities to map out new pathways for referrals and collaboration;

- Recruitment of key individuals (local 'product champions') across the partner organisations to take on ambassadorial roles in relation to their own stakeholders, e.g. GPs, Chamber of Commerce;

- Inter-organisation visits and 'shadowing' of colleagues to gain a better understanding of other people's work roles;

- Opportunities for staff from different partner organisations to work or train together, e.g. joint work with the client, such as employment coach and therapist meeting the client together, or joint training sessions;

- Regular facilitated opportunities for trouble shooting or rapid resolution of operational difficulties.

3.1.13   Sustainability and local development

Moving IAPT into main stream funding and the opportunities this offered both for improving and restructuring links with other parts of the service and

expanding the IAPT offering dominated the discussions on sustainability at the second phase of data collection.

At both sites there was commitment to establishing a single point of referral for all psychological services.

In Newham, commissioners originally felt disempowered by the centrally-commissioned IAPT initiative. As IAPT moved from demonstration to mainstream status, commissioners took the opportunity to review the service with the aim of reconfiguring the IAPT offering to bring in the third sector and to emphasise working in partnership with local independent community providers.

For Newham one of the main learning points to emerge in the transition from a demonstrator to a mainstream service concerned the failure of the IAPT and existing services to forge clear working pathways both between services and in the presentation of services to key stakeholders.

> *"…and, it was interesting, the conclusion of the review, really, because I mean there were certain very clear messages, issues over value for money, issues over duplication, principally, they're [IAPT and existing PCT services] targeting the same people and the obvious case that, you know, they operated as completely separate services, GPs could refer there, they could refer there, you know, where do they refer?".*

It was clear that this issue stemmed from the way the IAPT service had been developed over the course of the demonstration period. The re-design of the service around a single point of referral and the involvement of third sector organisations was seen as a way of tackling some of these issues.

The original aim at Newham was to involve a third sector provider in the single point of referral. It came as something of a frustration to the commissioners to find that they weren't able to use such a provider to directly employ IAPT trainees due to complications around conditions of employment. There remained however a strong commitment to greater integration between NHS and third sector services in the context of 'IAPT2'. Commissioners stated that they wished to see the IAPT model evolve, so that it was less of 'a service,' and more 'a system.' It was acknowledged that to increase third sector involvement in particular, a number of hurdles had to be overcome,

Other plans for the future at Newham include:

1. The mapping of existing provision and getting pathways in place

---

2. Greater third sector involvement across the partnership

3. Developing employment links

For Newham, one positive aspect of the demonstration period is the degree to which it has convinced commissioners about the significant role employment can play in the delivery of IAPT services:

> *"I was quite sceptical, originally, around how the employment bit fits into this, and you know, what are they delivering, but for a small service, their outcomes around helping people into employment have been pretty good, so,… I saw them as something that I wanted to grow, actually, and see that becoming more important, ..., and, ..., I wanted to not just think about employment, per se, ..., but, kind of, other, ..., perhaps slightly more generic support and advice that people could be given."*

At Doncaster the demonstration model was already close to that of the mainstream IAPT roll out, however, there were still lessons learned, particularly in terms of the management resource required to ensure the effective running of the service.

At this site, as for Newham, plans for the future are also concerned with developing a single point of referral and working to embed clinical pathways across the region:

> *"So there's now a mental health clinical pathway for [Yorkshire and Humber] which all PCTs are signed up to and are trying to implement … essentially some of that is about a single point of access for all Mental Health Services, rather than just single point of access for core of mental health, non crisis. So there's a – so in terms of the treatments, if you like, in terms of IAPT – so evidence based treatments for core of mental health problems, they're still being delivered, but there needs to be quite a bit of work done on what that front end of the service looks like"*

Other developments in Doncaster are concerned with the provision of specialist IAPT to specific populations. An IAPT worker has already been employed to provide a service to people with hearing difficulties and future plans are being developed to widen access to groups who may have previously been marginalised including ethnic minorities, and prison populations

Summary

This chapter aimed to describe the process of implementing IAPT from multiple organisational perspectives. It sought to assess the extent to which the demonstration sites achieved their implementation objectives, highlighting helpful and hindering organisational processes and issues.

More than 60 individuals participated in interviews (some at two timepoints). Interviewees were NHS employees and other stakeholders from a range of positions within organisations, providing perspectives at the individual, team, managerial and strategic levels.

Seven higher order themes emerged from the first round of data collection, reflecting the priorities of the set up period. These remained relevant at the second round of data collection and are used here to summarise and present findings. In addition 'Partnership working' was identified as an overarching theme which touched on virtually all aspects of the demonstration phase, and 'Sustainability and future plans' dominated discussions at the second stage of data collection.

The high profile, political nature of being a demonstration site, coupled with extremely short time frames for delivery made for extremely demanding conditions at both Doncaster and Newham.

The timeframe for the pilot had far-reaching consequences at many levels within the demonstration sites. At an operational level it militated against the adoption of best practice in the service set up, meaning implementation decisions were rushed, there was little time for consultation and the development of working relationships with partners and stakeholders, important mechanisms such as referral pathways at both sites and the IT system at Newham were not adequately tested prior to implementation and effective communication of the new service and engagement of stakeholders was limited. These problems in turn led to confusion or lack of shared vision, mistrust between professional groups and partner organisations and suspicion of the IAPT service. At an individual level the pressure to deliver and the consequences of rushed implementation created undesirable working conditions and conflict. Ultimately the limited timescale operated against the sites being a true demonstration of what could be achieved by the IAPT model and risked a poorer quality service to the user than could be expected if timescales allowed for good practice to be adopted at the development and implementation stage.

In terms of innovation, IAPT is a unique combination of:

- New service format

- New staff groups

- Highly flexible service

- Rapid access service

- Extensive partnership working with non NHS organisations

- A wide range of stakeholders & gatekeepers

- A service based on a proactive approach to care, with a remit for engaging with 'hard to reach' or non-traditional users of PTs

In addition, the role of assessment and monitoring in the IAPT service means it is both data and admin 'heavy'. The peripatetic nature of the service adds complexity to the level of data and admin support systems required. Getting admin or IT wrong impacted considerably on the operational capabilities of the service. In particular, The importance of a good IT system for managing the service and giving fast and accurate feedback about the service cannot be over-stated. This needs investment up front and time for staff to be trained in how to use it properly. If, as was the case for Newham, an IT system is being developed from scratch there needs to be realistic allocation of time and resource to do this and recognition of the delays this will cause to the effective operation of the service.

A key innovation in IAPT was to increase access by moving away from traditional service delivery methods and putting counsellors into a range of local/community settings. This aspiration was realised by both services, but posed considerable problems in terms of identifying and securing appropriate accommodation.

With regard to organisational systems, existing structures struggled to cope with the demands of the demonstration sites. This ranged from the timing of funding decisions and guidance on contractual issues for the new group of workers created by IAPT, to overall governance and accountability arrangements. As noted above, administrative and IT systems are crucial to the smooth running of IAPT and these were not always aligned to the IAPT model of service delivery and the peripatetic nature of the job at the start of the demonstration sites.

An important finding to emerge over the course of the demonstration period was a better understanding of the complexity of the management

role in setting up and embedding the appropriate systems to support IAPT delivery. Both demonstration sites drafted in additional senior management resource as the scale of the task became clearer.

Much time and resource was spent on the implementation process itself and many of the day to day challenges concerned the practicalities of grafting a new innovation on to existing services. Overall both sites felt they achieved their primary objective of demonstrating the viability of the IAPT principles of service delivery ie case management, brief intervention, a stepped approach and the use of telephones for rapid access. However, sustaining collaborative working with non-NHS organisations was difficult. Newham achieved a partial accommodation with a parallel employment coaching third sector service, Doncaster reverted more to a health service led rather than an inter-agency model.

Job characteristics, or the direct experiences of people working to deliver the new IAPT services was a story of extremes. On one level it was clear that case managers and therapists derived a great deal of personal satisfaction from their roles and where supervision worked it was highly valued and of great benefit to the overall service. Front line staff however, often bore the brunt of strategic or managerial failings as they had to deal day to day with challenging situations (such as inappropriate referrals or lack of supervisory support). These problems diminished as greater management resource was introduced to the demonstration sites.

Additionally, creating a strong IAPT programme changes the roles for others in adjacent services and teams; where these teams could not gain negotiated clarity about their roles, inter-professional tension was exacerbated. The notion of building provision into the service for resolving such issues was a frequent theme. Suggestions varied from joint site visits, joint working and shadowing to increase understanding and build working relations, to forums or cross partnership working groups with a trouble shooting remit. Whist such approaches can reduce clinical time, in the longer term they pay dividends in smoother operations and greater understanding between services. At both demonstration sites such approaches had been found helpful in building better collaboration.

Partnership working emerged as possibly the most challenging aspect of the IAPT delivery model primarily (but not solely) because of the timescale imposed for implementation. The time required for genuine partnership working is considerable, and it requires external drivers (e.g. commissioners), commitment at the top of the organisations, plus skilled

individuals in both the organisations needing to collaborate – these are the opinion leaders, change agents, product champions and 'boundary spanning' people in the Greenhalgh summary[10]. For the future, one of the most important messages for the expansion of IAPT (and other new services) is to operate to more realistic timeframes that support the adoption of good practice principles in the implementation phase. Working in genuine partnership with service users, a key aspiration of IAPT, remains a challenge and was not fully realised in either site.

Both demonstration sites have now moved into mainstream status and are planning for the future sustainability and development of the services. This has led to new and exciting developments in their IAPT services. Single point of access to mental health services is a key theme at both sites, as well as planning for how services can be developed to meet the needs of specific populations or communities. Work is also underway to build better and stronger partnerships and to link IAPT to other local and regional initiatives.

Both sites have demonstrated how an IAPT model of service delivery can be set up in spite of the challenging implementation timetable. Inevitably there have been set back and problems along the way, but there have also been great successes: At Doncaster the training and case management model integrated with the IT system, continuous monitoring of referrals and waiting times and fast response to problems is notable; at Newham the marketing drive to foster self-referral demonstrated what could be achieved in this area, equally the splitting of clinical and line management led to a better job experience for front line staff.

4 The patient experience of IAPT

Introduction

Patient experiences of engagement with a service are important to understanding how well a service is achieving its aims. Satisfaction measures look at patients' perceived reactions to treatment experiences and can influence and is influenced by treatment outcomes [14,15]. Currently there are no published reports of patient experiences of stepped care and IAPT interventions. We have considered patient experiences in two ways: through a brief satisfaction questionnaire and through semi-structured interviews. We report the questionnaire study first and then the interview study.

Patient Satisfaction Levels

4.1.1 Research Question and Design

The aim was to investigate patient satisfaction with the services patients received using data from the Client Satisfaction Questionnaire-8 (CSQ-8) [16].

4.1.2 Sample

There were two data sets available for analysis. As part of the comparison cohort study (data gathered from specially recruited cohorts in the demonstration sites and in comparable PCTs without an IAPT service; see Chapter 6 for details) patients completed the CSQ-8 at the end of treatment. In total 435 patients completed questionnaires.

The second data set comprised completed CSQ-8 from the Newham 'download data', which the service routinely collected as part of their clinical activity. Unfortunately, Doncaster did not routinely collect CSQ-8 data during the time frame of the study.

4.1.3 Measure

The CSQ-8 is a brief 8-item questionnaire assessing global satisfaction on a 1-4 scale, with a possible score of 8-32, and has good psychometric properties[15]. All written comments from the CSQ-8 were analyzed using thematic analysis [17].

4.1.4 Findings

Cohort Study CSQ-8

Overall patients from the demonstration and comparator sites showed satisfactory levels of satisfaction, M (sd) = 23.3 (6.1) and M (sd) = 22.2 (5.7) (maximum score possible = 32) respectively. Comparison between Newham (M = 22.3, sd = 6.2) and the comparator site of City and Hackney (M= 24.4, sd = 4.9) showed no statistically significant difference, t (80) = -1.76, p=0.08. Doncaster scores (M = 23.3, sd = 6.1) were significantly higher than those of the comparators, Wakefield (M = 21.3, sd = 6.1) and Barnsley (M = 21.5, sd = 5.2), t (326) = 2.65, p = 0.01.

However, 11% of Newham and Doncaster patients said that the quality of the IAPT service was 'poor' and 9-13% (Doncaster and Newham respectively) indicated that they were 'quite dissatisfied' with the service. Comments indicated that this was because the treatment they were offered was too short, or lacked continuity, and many found it difficult to engage with low intensity work, particularly telephone consultations. Some respondents indicated that they were still waiting to be seen and others had sought alternative help after receiving what they felt was a poor service.

Newham Service CSQ-8

Of the total number of referrals (n = 3371) a total of 1186 (35.2%) completed one CSQ-8 or more (range 1-7). Of those discharged from the service (n = 2811) the proportion of those completing one or measure is similar (n = 996, 35.4%). The majority of these (529, 53.1%) completed one measure only. The last measure completed was used for this analysis. As the CSQ-8 scoring system uses the total score, only those who completed all 8 items were included (n = 1039).

The mean CSQ-8 score for all referrals was 26.98 (sd = 4.13). Of the 1039 patients who completed a measure, the majority (n = 748, 72.0%) had one or more intervention sessions. Those who had one or more intervention session (M = 27.28, sd = 4.05) were significantly more satisfied than those who had an assessment only (M = 24.91, sd = 4.39), t = 5.77; p < .001.

The majority of clients completing a measure were those who completed or agreed to end a course of treatment (83.2% of all those completing treatment) compared with 11.3% of those who did not complete. When 'completers' are compared to 'non-completers' (at any stage of treatment dropped out/terminated early, were unsuitable or ineligible, referred on to another

service or did not need the service) 'completers' (M = 29.31, sd = 3.65) were significantly more satisfied than 'non-completers' (M = 25.15, sd = 4.14), t (862) = 11.96, p < .001.

Those who had a High Intensity intervention were more likely to complete a satisfaction questionnaire; 76.7% who had a High Intensity intervention only completed a CSQ-8, compared with 56.3% who had a Low Intensity intervention only. A one way ANOVA with Bonferroni adjustment for multiple comparisons showed that those who had a High Intensity intervention only (M = 27.71, sd = 3.94) were significantly more satisfied than those who had a Low Intensity intervention only (M = 26.34, sd = 4.24) and those who had both interventions (M = 26.89, sd = 4.14), f = 13, p < .001. Those whose intervention was by telephone only (M = 25.90, sd = 4.10) were less satisfied than those who has some face to face intervention (M = 27.66, sd = 3.97), f = 7.17, p < .001. Those who were receiving benefits (JSA, IS, IB) (M =26.35, sd = 4.48) were less satisfied than those who were not (employed, student, retired, homemaker) (M = 27.37, sd = 3.93), t = 3.35, p = .001.

There were no differences in levels of satisfaction between men and women, age, different ethnic groups, for those who self refer compared to those who are referred by their GP or other professionals, or by whether the primary diagnosis was a mood, anxiety or other disorder, all ps > .05.

Patient Experiences

4.1.5 Research Question and Design

The aim was to provide detailed analysis of the pathways patients took and their experiences within the demonstration sites. The following questions were considered through analysis of interview data,

What are
- Patients' experiences at different stages of the services?
- Patients' experiences of changes in work status?
- The experiences of patients who did not complete treatment?
- The experiences of patients who had good and those who had poor outcomes?
- Patients' experiences of 'stepping up'?

These questions were investigated through a qualitative exploration of key themes using in-depth interviews and analyses informed by the Framework method.

4.1.6 Sample

The interview sample comprised patients discharged from the demonstration site IAPT services. In the proposal we stated the sample would comprise a range of participants in terms of socio-demographic characteristics and levels of functioning and symptoms and participants with good outcomes including those returning to work and those with poor outcomes and service experiences. We intended to purposively select participants from the cohort sample. Informed consent would already have been obtained and we would have data from which to purposively select participants. Because of delays in beginning the cohort study, we obtained NHS ethical agreement to seek participants from any patient who had been through the IAPT services. Until patients consented to be part of this study we could not access any demographic or treatment outcome details, and therefore purposive sampling was not feasible. Instead, we interviewed all from whom we had consent. We then obtained details about our sample from services with a decision to continue interviewing until we had a sample that included patients from different socio-demographic groups, with varying intake and outcome scores on the PHQ as stated in the proposal. Our first check of our interviewed sample showed we had achieved our aims. Table 10 shows the sample characteristics and, to put these in context, the case profile of the service as a whole at this period.

Letters were sent to 350 patients in Doncaster and 437 patients in Newham, beginning with the most recently discharged. In addition, 100 patients at Doncaster who had been 'stepped up' and who had been discharged following CBT or were coming to the end of treatment as part of the IAPT service were sent invitation letters. From this combined sample 54 patients (including 11 patients who had been stepped up) from Doncaster and 23 patients from Newham were interviewed.

Of the 77 patients interviewed 25 were men, their ages ranged from 16-69 and 14 were of non-white ethnicity. At intake in to the IAPT service 33 patients were in the mild to moderate range on the PHQ and 44 in the moderate to severe range. At outcome 40 patients had returned PHQ scores that indicated 'recovery', 24 had not reached 'recovery' scores, and for 7 their scores had deteriorated (6 did not provide end of treatment scores).

Project 08/1610/154

Table 10. Sample Characteristics

	Qualitative Interview Sample		Overall Service Case Profile (1/4/06 – 30/4/09)	
	Doncaster n (%)	Newham n (%)	Doncaster %	Newham %
Gender				
Male	16 (30%)	9 (39%)	35%	39%
Female	38 (70%)	14 (61%)	65%	61%
Age				
Mean	45.15	36.78	38.2	37.7
Range	16-69	23-56	16- 89	14 - 87
Ethnicity				
White	48 (89%)	9 (39%)	99%	48%
Non-white	0	14 (61%)	<1%	52%
Not known	6 (11%)	0		
Primary referred diagnosis				
Anxiety	1 (-)	4 (18%)	Mood disorder 48%	Mood disorder 44%
Depression	26 (49%)	9 (41%)		
Mixed Anxiety Depression	21 (40%)	4 (18%)	Anxiety disorder 49%	Anxiety disorder 48%
Not known	5 (9%)	5 (22%)		
One or more High Intensity Session				
Yes	18 (33%)	12 (52%)	6%	57%
No	36 (66%)	11 (48%)	94%	43%

4.1.7 Procedure

Individuals eligible for inclusion were: anyone discharged from the Doncaster or Newham IAPT service in the reference period (October 2006- September 2008), starting with the most recently discharged and working backwards in blocks of 50-200 people (except for the 'stepped up' sample, see above).

The term 'discharged' – refers to early leavers, completers and people who had no contact with the service following referral; it included those who were 'stepped on' to counselling straight from initial assessment, sent to the CMHT, or signposted to other services, for example, CRUSE, Alcohol advice etc; finally, it also includes those whose last treatment contact was some considerable time before their official system 'discharge' date.

Recruitment was by written invitation, from the service to the former patient, with freepost envelopes enclosed for replies. In this way, individuals had to directly volunteer their contact details to the University of Sheffield, and their identity was protected until they chose to do so.

Initially, recruitment was via a single invitation letter [signed by a Case Manager or therapist] enclosing a full length information sheet and consent form from the University of Sheffield to be returned by post. Subsequently, this was amended to a two-stage recruitment process:
- An initial covering letter from the Case Manager/ Therapist enclosing a short 'flyer' from the University of Sheffield inviting expressions of interest and asking for telephone contact details. Once these details had been received, this was
- Followed by telephone contact (by University of Sheffield researchers) to explain the research, and then a mail out of the full information sheet and a consent form, together with confirmation of interview arrangements (see Appendix 4.1).

In addition nine participants were included in the study via the cohort recruitment process (see Chapter 6); these were individuals who wrote comments on their questionnaire returns indicating that they would be keen to be interviewed and/ or willing to answer further questions. These were contacted by email initially to confirm their interest, and were then telephoned and sent an invitation letter, as above.

All those interested in being interviewed were offered a mutually convenient appointment time at a location of their choice (e.g. GP surgery, community venue, or in their own home). Appointments were confirmed in writing and

again by telephone on the day. Consent forms (sent in advance) were discussed and completed in person at the time of interview.

4.1.8 Recruitment of non-English speakers (Newham only)

Newham's Service Administrator, who spoke 3 languages (Punjabi, Gujerati and Urdu), volunteered to contact by phone patients who had received therapy in one of these languages. The Newham IAPT Service database was searched for referrals between Jan-Mar 2009 of patients who had received treatment in Punjabi, Gujerati or Urdu and the Service Administrator attempted to contact these individuals. With the patients' agreement, their contact details were then sent to the University interviewers to arrange an interview (n=4). The interviewers were three trainee clinical psychologists fluent in the above languages who volunteered to be part of the project team. Four potential participants were contacted, two interviews were arranged, and one interview completed.

4.1.9 Interviews

The interviews were conducted using a semi-structured topic guide (Appendix 4.2). A draft topic guide was drawn up in discussion with service user members of the research team. The guide covered the following key areas: acceptability of the service and treatment pathway, service experiences, experiences of change, roles and responsibilities and views of staff. The key purpose of the interviews was to capture the service user's experience from illness/ symptom onset through to current circumstances, covering key aspects of their service referral, contact experience and perceptions of change/recovery. The majority of respondents were also asked about their work status and the relationship between their health and any employment.

The interview topic guide for those who were stepped up was modified and greater emphasis was given to asking about experiences of 'stepping up' from low-intensity to high-intensity CBT treatment. The topics covered in these interviews included: initial impressions of the IAPT service; experiences of case management; experiences of stepping up; and experiences of CBT (see Appendix 4.2).

All of the interviews were conducted flexibly; ensuring patients were comfortable with the interview situation and were able to tell their experiences. It was made clear to patients at the beginning of the interview that they were free to stop the interview at any time and to withdraw their consent. At the end of the interview patients were thanked and asked if there was anything

else they wished to add about their experiences with the service or about the interview itself. The interviews were between one and one and a half hours long. There were five interviewers (GH, RH, KD, JC, RHo). The interviews were digitally recorded with the participant's consent. Digital recordings were stored securely and transcribed and anonymised by University of Sheffield staff.

4.1.10   Framework Method

The analyses of the interview transcripts were based on the Framework method [18]. This method allowed the evaluation of questions and themes that were part of our research questions to be addressed as well as incorporating an inductive element that highlighted the experiences of those receiving IAPT services. The analyses involved five stages: familiarisation; identifying a thematic framework; indexing; charting; and interpretation.

The transcripts (and fieldwork and relevant researcher diary notes) were summarised against each of 6 high-level codes, and 29 sub-codes and entered in to an Excel spreadsheet by four analysts. The purpose of this was to enable the data to be read thematically, and to keep individual cases/ stories whole. Sections of the transcript where more than one code applied were only charted once, and cross-references were used to ensure all relevant information for each code was signposted. In summarizing the transcript text, analysts sought to keep as much of the respondent's original speech style as possible, and included links to enable the summaries to be traced back to the relevant section of the transcript, and to locate quotations.

Once all the transcripts were charted in this way, data on service use and demographics was added in from the IAPT services' management information systems. The combined datasets within this Excel spreadsheet then became the primary resource (the 'frame') for the next stage of analysis. The final coding framework is shown in Appendix 4.3.

The first analysis carried out was thematic. This followed three key steps:
- For each code in the frame, an initial detection exercise was carried out to highlight key points; these were then extracted on to a set of summary sheets for each code.
- The extracts for each code were then grouped together in to meaningful conceptual categories to form a second summary sheet.
- Finally, the categories were examined to create higher order categories, and to assess the relationship between categories – and a classification of the key points made.

At the same time, a second type of analysis was carried out, looking primarily at differences between cases. The purpose of this was to capture the specific experiences of three sub-groups of patients: those who did not complete treatment, those who had poor outcomes, and those who were 'stepped up' to high intensity work.

Finally, where appropriate, the thematic analysis was extended to include some typological analysis of cases – for example, identifying groups according to their relationship to paid employment, or previous service use. Each analyst took responsibility for part of these three types of analysis. Regular meetings were held to discuss the emerging findings.

4.1.11   IPA Method

The 11 interviews conducted with people who had been stepped up, in addition to the Framework analysis, were analyzed using Interpretative Phenomenological Analysis (IPA) [19] as this method seeks to explore participants' understanding of the stepping up process. Analysis followed standard IPA procedures and the transcripts were read a number of times when significant concepts or ideas were noted. Salient themes were then identified and checked back to the original transcript until a list of themes for each transcript was developed. These lists were then consolidated into a list of master themes.

In addition to the main researcher, three transcripts were analysed by a service user researcher who developed emergent themes, which were then discussed with the primary researcher.

This project formed part of Rachel Horn's DClin Psy thesis (Horn, R., 2009). A summary of the findings from this project is presented in the section on stepping up.

Findings

The findings below are summarised according to the approximate sequence of patients' journeys through the services, including their views on the experience at each stage. There then follows analyses of issues specific to particular sub-groups of patients. The site, gender and participant number of each respondent is shown in brackets after each quotation (D = Doncaster, N = Newham, F = Female, M = Male, n = respondent number).

4.1.12   Seeking Help

Patients tended not to seek help until they had reached a point where they could no longer cope. This was conveyed in a variety of ways; *'I just lost it'*; '*everything came crashing d*own'.

> *I was just going to work every day and crying and I just couldn't cope and then one day it were just like the final straw. (DF1)*

This was sometimes brought on by a culmination of a number of stressful events or a stressful experience triggering more long-standing problems. Where the participant identified a main cause, this tended to be work related, bereavement, relationship problems or diagnosis of a serious health problem. Some stated either they could not put their finger on what had caused their problems, or initially, they did not recognize that the physical symptoms they were experiencing had a psychological component. In these cases, conditions were either identified by their GP, or the individuals were persuaded to attend the IAPT service by close relatives or work colleagues.

> *I'd not really gone to the doctor's thinking I was depressed anyway this time. It was sort of, you know when there are lots of non-specific symptoms and you just generally feel run down and he sort of said to me 'Do you think you might be depressed?' I thought ... now you mention it yeah possibly. (DF2)*

4.1.13   Accessing the Service

Participants experienced a variety of referral routes to the IAPT service, primarily through their GP but also via the practice counsellor, self-referral, and employment services. Some expressed confusion about where and why they were referred.

> *There was so much going on at the time, what it is I am also on Diazepam – due to my problems and that, so I can't really remember what it is like – it is confusing - So I can't remember what the interview was about, but I know I remember going down there and the thing was … I don't know, 'what am I doing here? , I don't know who referred me, no not at all. (NM1)*

The option to self-refer was generally received in a positive light. Participants mentioned finding out about the service through leaflets advertising the service in the library and a local publication posted through the door, or were given the number by their GP. Those who discovered the service via local advertising liked that they had the option of contacting the service direct rather than

having to go through their GP. The feeling of having some sense of control over where and when referred was evident.

It's a new magazine … I kept it aside, you know, it's quite an unusual thing that you can actually just ring up people and say excuse me I'm going through difficulties, but usually you'd have to go like through a GP. So, I mean, I had a crisis in my life at the time … I remembered that piece of paper so I went and found it somewhere in the cupboard and sort of called somebody up and that's how it sort of started, if you like. (NF1)

4.1.14   Treatment Choice

For those referred by their GP fast access to an alternative to medication seemed to give some people a sense of having a choice.

GP suggested IAPT would probably be a good idea … so I just thought I'd give it a go before any type of medication, I think it's so easy nowadays for somebody to just like give you a tablet and say 'oh you just take this' and I think it sometimes it's better to try other routes before you go down that one. (DF1)

However, many patients spoke about the lack of information and lack of choice provided by the GP. For some, little information was given about what the treatment would involve, what type of service they were being referred to, or whether any alternatives were available to the IAPT service.

I don't think they actually explained the process to me, and as in why it would be beneficial for me, you know, right, why that therapy rather than another therapy, and that sort of stuff. (NF2)

Lack of awareness and understanding was one factor that participants identified as making it difficult to raise any issues they might have.

I think it would have been helpful for me to know what kind of service I was being referred to and for me to make an assessment really of whether I felt that it would have been… (DF6)

The state of mind of the person attending the GP sometimes affected how information was understood, for example one patient stated that the options were not discussed because she was 'crying all the time'. Sometimes, patients let the GP recommend what should happen, either because they trusted their opinion, because they didn't feel that they were in a state of mind where they could make a decision, or because they were so desperate to get better that they were prepared to acquiesce and to try anything that was offered.

The only alternative they gave me was this service … the only reason I took it was firstly to see if it does make a difference, and secondly because I wasn't going to receive any other treatment otherwise. (NM2)

The GP said 'I think you would benefit from talking to somebody' and that's when he said about this, the IAPT scheme, and would I be prepared to, and I said yes, and at that point I were … prepared to try anything to make me feel better. (DF3)

The opinions of some patients however appeared either not to be respected or to be ignored, by a refusal by the GP to give them their preferred option. For example,

I had a new doctor and she wouldn't refer me. Why? I didn't actually ask why…. When you're feeling depressed…. Your heart beats and you think what's the point you know. (NF3)

4.1.15   Waiting Times

Most participants reported that they were contacted by the service within a few weeks, which was appreciated.

I was seen quite quickly which made a lot of difference because at that moment in time I couldn't deal with things, it was frightening (DF20)

However, waiting time tended to be longer for High Intensity work or when people were stepped up, which appeared to lead to dissatisfaction.

Well my GP gave me, um, a phone number to do a self referral, which, I did and then I still had to wait, sort of, about six months before I even got to see someone, so six months of total isolation. (NM3)

When there was a waiting list some participants were sent details of how to access self help packages. Experiences of this differed from a perception of receiving no help to feelings of being kept informed of the current waiting list situation. There was also dissatisfaction expressed where the person had prior knowledge of the IAPT service and had higher expectations relating to access and the type of service they would receive.

4.1.16   Initial Contact

The first meeting with a practitioner was often important in engaging patients and providing hope, *'it felt like there was somebody there who was going to offer something'.* Sometimes, however, patients expressed disappointment

because they had not expected the treatment plan to consist of phone or computer work instead of face-to-face meetings or the first meeting felt impersonal.

I remember I wasn't very well at all. I didn't find, I found it a bit, because it was very what I call 'clip-boardy' …. and I was there having a panic attack and I found it a bit off-putting. I know it had to be done, obviously they need to know what they need to do but it was very clinical. Erm, not very, there didn't seem to be a lot of caring bits in it, as it was your first one. (DF4)

4.1.17   Choice within IAPT treatment

Patients valued being offered choice by the IAPT services, even if the choice was minimal. For example, patients were offered the choice of a book or Beating the Blues (BtB, a computer-based intervention for the management of depression). However, some patients felt they didn't have a choice. Any decision as to the appropriateness of what they were offered was taken out of their hands or the decision appeared to have already been made for them.

[The practitioner] decided that we were going to do the BtB thing and I'd do one session at home and then one session at the centre with him … I wasn't aware of any other type of, you know, like therapy or anything, I was just told I was doing that. (DF1)

Sometimes the least preferred option was offered in such a way that the patient acquiesced.

I did tell the case manager that I didn't like the telephone and she said she would put me on the waiting list for counselling, but I've not heard anything. I felt uncomfortable with telephone but gave it a go; but it wasn't what I'd expected. (DM1)

Patients spoke of feeling vulnerable and their need for reassurance when attending the IAPT service. Patients also spoke of the need to be self-confident and assertive. Those who were convinced of the value of their treatment appeared to fully engage and find benefit, such as working through some of the exercises and the booklets. For some this determination also related to getting referred in the first place.

It's only through sheer persistence that I've got to see anyone. (DF5)

Some patients found it difficult to be 'strong and demanding' in relation to their treatment choice. This resulted in not getting the help they wanted, which sometimes led to them blaming themselves for the intervention not working.

I was trying to talk about things that I desperately wanted to talk about but I couldn't get them in the [telephone call], I just couldn't get those issues in... I felt very unsafe...I think it is asking a lot to find a time and a place where.. [you can] speak to someone about things that are quite personal...you either try to look after yourself, or you are very demanding...I tried to make sure that I didn't expect too much, or expect something that wasn't on offer...when I look back, in that first session, I should have said 'no, I don't want to do that', I should have been a bit more assertive but I thought I would give it a go...because there seemed to be nothing else other than that... It is quite a big thing to ask for help in the first place and for it not to work is an even bigger bump. (DM1)

4.1.18   Telephone work

The most common Low Intensity interventions following a preliminary face-to-face meeting were weekly or fortnightly telephone calls, supported by use of information booklets, written exercises, and homework or BtB.

Reactions to telephone work varied. For some it was preferred to face-to-face meetings because of the comfort and flexibility it offered. Some found telephone work a positive advantage because it offered greater convenience and felt safer.

I were a lot more relaxed talking to him over the phone than what I would have been if I'd been down to the surgery. (DM2)

But most saw telephone work as second best; they felt less connected to the Psychological Wellbeing Practitioner (PWP), found it easier to avoid talking about difficult topics and difficult to find privacy for the phone call, or they could hear background noises from the PWP's end. For some patients it appeared easier to not tell the PWP how they were really feeling because of the indirect nature of the communication.

I did find telephone interviews quite distressing in a way, I dreaded them when they were going to happen and didn't really feel the benefit or the value of it because it wasn't face to face. I just found that it was easy to be distracted and to say what she wanted me to say, not that she ever led me into saying anything but just, it was just easy to answer the questions quickly to get her off the phone and then go back to what I was doing ... (DF7)

For some there were practical difficulties with telephone work.

We would make an arrangement for a time for her to ring – err – which was OK at first because I was actually at home, I was on sick leave for 2 months and erm – and - and so that was accessible but then when I started to feel better and I went back to work then it became a little bit more difficult because they had to fit it in at work and you are not in an environment where you can go and sit and have a conversation about how you feel. (DF6)

Some patients appeared to have a negative experience of therapy delivered over the telephone, which could be due to the mode of delivery or a negative relationship with the therapist.

Telephone therapy – that was a failure because I'm no good on the phone erm, a lot of people aren't, I needed to see somebody face to face, and the lady seemed more to concentrate on my diet erm, and also because I told her that I wasn't very good with cooking and I tended to skip meals, and this was all down to the anxiety I couldn't eat half the time but she kept ..... it was exposure therapy over the phone, and she was pushing, and pushing, and pushing, and she couldn't understand how it took me all my time to get in the car to go to work – she was telling me to stay in shops for 2 hours and I wasn't ready for it you know, so that wasn't for me... [It was a] terrible waste of time. (DF8)

4.1.19   Low Intensity Interventions

For many patients the structured approach of low intensity work was valued. Patients reported liking the way problems could get broken down in to smaller parts, by writing things down & looking at things together with the PWP. They found it helpful to work through the booklets section by section, reviewing homework exercises, and having clear targets. Some found the booklets helpful in explaining specific conditions, such as anxiety and avoidance, and liked the format,

She used to give me…a couple of booklets to do – [these] were quite helpful…they were just mainly scenarios and quizzes – you know – to get you thinking, as well as… – ways of improving yourself, …and if you keep doing something it becomes like a second nature to yourself. So if you read that book probably about 5 – 6 times you would take the input from it and hopefully you should cause a change within yourself. (NM2)

Others found the interventions were not pitched at the right level – they were either regarded as too simple and superficial, or conversely, too complex and

hard to do, if unwell. Patients also expressed concerns about interventions being used without sufficient support.

I'm quite a verbal person anyway, anybody who talks to me, you know, they can, sort of, bring out things by just encouraging words and statements and mannerisms and, you know, smile, you know, anything, raising an eyebrow, who cares! The thing is, but writing out a diary, for me, just didn't work...I just looked and I thought, I can't be bothered, couldn't do it.... I mean, now, I would just, jotted some, a few things down and it would have been, not hard. But then, if you're feeling really, really ill, you just don't want to do it. (DM3)

4.1.20   Computerised Self-management Packages

The use of computerised self-management packages, such as BtB, was liked by some and not by others.

I found it nice being on the end of the computer and not to be embarrassed by what I might type in.(NF4)

BtB was a bit dry compared to talking therapy. (NF5)

Some liked the way they could fit it into their daily routine.

I was quite happy with using the computer program, it was easier to fit it round you know, … I didn't have to rely on other people to babysit while I was doing it erm, that kind of thing you know, it was something I could just kinda do. (DF9)

The extent to which it was supported by contact with the PWP, how it was introduced and if it was introduced once rapport with the PWP had been established helped patients engage with the intervention.

I found BtB very helpful, once I engaged. I knew what I wanted from it and I could go back now. At first it was difficult because no one talked about what happens when you have chronic pain--if someone had shown me earlier it might have helped. (NF4)

For some it was hard to find uninterrupted time to go online, or they had technical problems.

I weren't impressed with it at all. Erm, it kept kicking me out. ... I'd only get so far and then it you know the pages would just go blank and or it would freeze. (DF10)

4.1.21 Contact with PWP

Patients talked about factors they thought contributed to the quality of their contact with the Low Intensity service. They valued practitioner flexibility in scheduling appointments, but were also acutely aware of the heavy workload of staff.

...even though we had an allocated hour, it could be slightly less, but for most of the time slightly more. She was flexible.(NM3)

I understand why they use the phone because if they had to drive everywhere they wouldn't be able to fit everybody in. (DF4)

Patients valued the continuity and consistency of contact with a PWP. They also valued the strong commitment of the PWP, which was noted by the ease of getting hold of them, that telephone appointments were held punctually, and kept to time and that when a patient did not respond to a telephone call, this was followed up on. Patients indicated that being called in for reviews and feedback on how they were doing was also useful.

It was good to have someone there to ring up and say you haven't completed it, are you having any problems or is there something you wanna talk about? You know, just to be able to say well I don't really understand it or no I haven't really got around to it or just – and then make sure you do it you know, sort of – so that kept you – that was good.... (NF4)

I think after so many sessions on the phone she arranged, she said, you know, we have so many sessions on the phone…. and the then you can come in to review how you're getting on…. We sort of then had like a review and we went over everything and….how I was getting on and if there's, you know, any improvements or if there's things I'm still not sort of happy with. And then she referred me then again to somebody else after that. (NF6)

The converse was true for some patients, who complained about the lack of follow up on homework, on missed appointments, and/ or after ending treatment. This included not being able to get hold of the PWP, so for example sometimes telephone calls not returned or non-response not checked up on and the fact that sickness absences were quite common and that staff turnover was at times an issue.

I asked her about the booklet and she said, "don't worry about it". So I was asking her, now why is it that they give me a booklet to read and we are supposed to discuss it, I have gone through it bit by bit and it is like nothing has come out of it. (DM4)

4.1.22   High Intensity sessions

Patients appreciated regular meetings with their therapist. Many praised the collaborative approach; said they felt understood, trusted their therapist and welcomed a chance to talk through issues and look at things differently.

Suddenly there was someone there who could make me feel better after speaking about it. (NM4)

Some patients described their therapist as being like a 'mentor or a guide' and some spoke of how their therapist helped find a focus or goals to work on.

There were 3 goals….and achieved these by breaking them down into meaningful action. (NF11)

Some patients thought that they improved much quicker through CBT sessions than they would have through telephone work, and for some they understood their problems better through a CBT explanation.

The lady gave me my life back. She made me understand things and I started to realise that I could beat it... The CBT gave my problems a name and explanation, which I didn't get from the case manager and computer. (DF8)

4.1.23   CBT Interventions

Patients found the following High Intensity techniques helpful: methods to think positively; using flip charts; experiments; homework tasks and diaries*.*

One day when I was feeling really low…[he] got out some flip charts and put the negatives down one side and what I had done with my life on the other. There was more what I had done and straight away I thought he was really good (DF26)

Patients also valued discussing which intervention might be useful.

The, er, CBT screening was very good... it was more, er, therapy focused. And looking at different types of therapy which might be, er useful for me..(DF27).

Patients described some of the benefits they experienced with CBT that included feeling more calm, confident and better able to cope.

It gave my confidence a boost and helped put things in perspective and helped me deal with things bit by bit (NF1)

However, some patients thought that their therapy was too short or that they did not have the opportunity to deal with some of the strong emotions they were experiencing.

I felt like I had this huge well of emotion inside me and with the CBT if you looked like you might cry, he sort of nipped it in the bud and it is like you have got to control your feelings and think positive thoughts and alter the mood and alter the feeling. I kind of needed to feel this and wanted to feel this and express it with someone and I wasn't given the chance to do that... (DF13)

4.1.24 Outcome Measures

The use of the outcome questionnaires received a very mixed response, with some patients finding them very useful as a way of monitoring their progress, and others feeling discomfort at the experience, disliking some of the questions, and finding their answers looked more positive than they felt. Some patients reported providing inaccurate answers in order to 'please' their PWP, or complained that the questionnaires took up too much time, and were for the PWP's benefit rather than their own. Those who appreciated using the scores themselves for self-monitoring purposes, also tended to report liking the reviews/scales in BtB.

When you're speaking to your counsellor, your case manager, there is a pull on you….for instance, you could be working on these things, you know… when you fill in your chart at the end …..I just felt a pull … I knew she wanted me to improve…. it wasn't objective… it was just something not right about that. When you get off the phone I think, I think I put a better spin on that than actually I'm really feeling. (DF11)

You have to do the questions over the phone ….so, technically if you wanted to you could make it up anyway, if you wanted to fiddle the system I suppose, but eventually it will get sent somewhere else.... apparently…the case manager has an interview with her manager, and they go through each person and they discuss why the scores are not going down, and after a certain amount of time if they're not able to erm reduce your scores then that

was it, they think you need some other form of more intense treatment. (DF4)

4.1.25   Relationships with IAPT PWP and High Intensity staff

Patients talked about the importance of their relationship with their worker, whatever intervention they were receiving. Trust, listening, encouragement and a belief in the patient's capacity to recover were all important aspects of the relationship.

A nice person, into her job, enthusiastic about it, you can tell, and you know wanting people to recover. (DF5)

They have all been friendly and they have been professional but friendly meaning to talk to. You can approach them; I have been able to approach them all. They have not made me feel stupid, if I didn't feel, well I felt like if something happened I was in safe hands. (DF14)

It was important to patients that their IAPT worker treated them with respect, enabled them to feel safe when they were at their most vulnerable, and understood them.

She showed respect and dignity; she was brilliant and always on time. (DM5)

It was comforting to know that there was someone there that understood and I just put myself in her hands. (DF15)

Positive comments also related to the workers' perceived good communication and interpersonal skills.

Picks up on your eyes, expression, body language. (NM3)

CBT therapist had a great sense of humour and she was sympathetic but not patronizing. (DF8)

Where there were negative aspects to the relationship, comments tended to reflect a perceived lack or absence of some of the positive qualities cited above, or a fear of saying the wrong thing.

I didn't know how to relate to him--I was inhibited and didn't think that it was for me. He was really nice, but I worried whether I was saying the right thing. (NF4)

For others there was a mismatch between their expectations regarding age and experience of the worker.

She was maybe younger age group to me - talking to me in textbook 'this is what it should be like', or 'this is what you should be doing', but I wasn't able to do those things. (DF7)

A number of obstacles to developing a closer therapeutic relationship were reported. These included insufficient time, poor interpersonal and communication skills, and a lack of experience.

I struggled to relate to her a little bit. I did feel that in what she was doing had been some rushed training, I didn't feel like I could put confidence in her, there was not that experience there and that could have been me reading into [the] thing, I don't know. (DF13)

I am still unaware of the 'qualifications' of the person I was speaking to. (DF11)

Patients appeared to be particularly dismissive where they felt their workers' responses were not responsive to the patient's needs and were formulaic or impersonal, or the service left them without information when their IAPT worker was off sick or had left the service.

You could have got a robot to do [it]. There wasn't really any feedback and it just seemed a bit impersonal I think whereas.... someone with more experience might have been able to use more instinct and sort of bounce off things better and relate to you a little bit more I think.(DF13)

Several patients reported being seen by more than one IAPT worker without this signifying a 'step up' (or 'step down') in their treatment, and were able to contrast their different experiences.

My first therapist... he was very sort of calm and reassuring …. he seemed like the ideal therapist....[Then] I had this other guy and he er – he was very different....he just went 'why?' [laughter] when I said something like.... and first I was like 'well now you're acting as if I am actually crazy' (NF7)

4.1.26 Relationships with GPs

GPs also played an important part for many patients in their treatment experience and some were able to offer additional services, for example referral to council exercise facilities or a mental health nurse. They also offered practical advice and suggestions. During and after IAPT treatment many patients continued to maintain contact with their GPs. For some it was a case of letting the GP know how they were getting on, about their thoughts on

returning to work, and/ or to review medication in light of how they were doing or following on from discussion with their IAPT PWP.

She's been asking me how I feel and what treatments and like if I've felt any better by seeing different people and she's advised me not to go back to my place of work and then she's been giving me sick notes. (DF1)

For others it was to raise concerns about their IAPT treatment and/ or get a 'second opinion'.

I thought it was important to let my GP know that the service was not helping me--if I hadn't contacted him he might have thought everything was OK. I could have fallen through the net. (DF16)

Many patients were grateful that their GP was prepared to listen and be available.

The GP was fantastic. Listened when started going into depression just before my mother died and noticed I was self-harming. I had an appointment to see her every day. (NM4)

Others were less satisfied, or unhappy about the lack of continuity of GP care.

I was left in like a situation where the GP refused to – erm – give me any more treatment or anything unless I took up with this – the choices which I had, one was to take medicine, because I was refusing to take medicine for a while, and the other service was [this one]. (NM2)

4.1.27   Improvements to the Service

When asked about what improvements they would like to see patients suggested a number of things. For some, the service needed to be better advertised and more GP involvement encouraged; for others, the service needed to offer more flexibility, for example, a more flexible mix of telephone and face to face meetings; better communication was also mentioned, particularly where there was a break in continuity of treatment due to staff absences. Other suggestions included: provision of an opportunity to meet others in similar positions; help with travel arrangements; longer duration treatment; greater continuity of care; and help in preventing relapse.

The main improvement I'd like to see is just not giving you a time limit of like sort of twelve sessions…I think if you want people to really get better… I worked at it so hard, that erm to cut somebody off who is really trying to get better, is false economy really. (DF11)

You need more support… and you need group sessions I think…. it would be lovely to meet other people in similar circumstances… because mental problems absolutely sap your energy… Say somebody was experiencing the same as you, you could all say 'right, well tomorrow or one day this week I'll come round to your house and we'll do your kitchen or whatever. (DF12)

4.1.28   Informal Sources of Help

Some participants spoke positively about informal sources of help with family and other important relationships being key,

Most important thing is knowing my boyfriend actually is really really supportive and you know he kind of understands what I'm going through (DF17)

Mixed views were expressed about informal sources of help. Some participants experienced their family and friends as supportive or as a useful distraction from their own problems, whereas others did not receive family support.

Some people, like my family they have never understood why – what counselling was about and I don't really discuss my personal issues with them anyway but they wouldn't have understood. (NF7)

I had thoughts of wanting to die erm, it were my children that stopped me thinking that just – I felt really sorry for myself, just completely and utterly...just a completely self absorbed (DF18)

4.1.29   Other Services Accessed

Some patients had had previous experience of a talking treatment, complementary therapy or support service (e.g. Victim Support, CAB). Types of help received included counselling via the GP practice or workplace, contact with a complementary therapist or with secondary mental health services, such as a mental health nurse, or psychiatry, clinical psychology or other specialist treatment.

The experience of other treatments or contacts seemed to make respondents more discerning in their appraisal of IAPT. For some, IAPT was an improvement, because of speed of access and type of help offered. For others though, it was less good than what they had experienced elsewhere; the main complaints related to its time limited nature, being too 'clerical' or 'superficial', not giving enough face to face contact, and not allowing for enough follow-up.

High intensity treatment was generally rated more highly than low intensity by these patients.

> *Most useful was the CBT therapist. That's most useful, because she made, she made change emotions a bit… It has helped me do things it has. (NM5)*

Some individuals received counselling from another source in parallel with their IAPT treatment. In some cases, this seems to have been a response to frustration about having little contact with the service due to a PWP leaving or having time off sick. In other cases, IAPT seems to have prompted a search for a more in-depth talking treatment e.g. through a voluntary sector provider.

> *… The IAPT counselling helped me to deal with the here and now, but the symptoms such as obsessive compulsive or you know mind reading, they did a booklet on and stuff like that…. And it's built your confidence, but then the other counselling is much more in depth and erm helps you to deal with, helps you to talk about big issues in your life that you've maybe not dealt with. (DF11)*

Respondents who had had no prior experience of a talking treatment tended to rate their IAPT experience very highly, sometimes as a transformative experience. Some of these still commented on its briefness however, and not everyone was confident about returning for further help – either directly, or via the GP.

> *You know it was smashing, but a lot of my friends feel that I have pulled myself through it but I don't think I have.(DF19)*

> *They only give you a quota, they give you a certain number of sessions and that's it, then they wash their hands of you. There's no feedback, no follow-up. (DM2)*

4.1.30   Ending Treatment, Moving On

Patients also spoke of the importance of choice and negotiation at the end of treatment whether from Low or High Intensity work. For some the ending was positive. Many felt that treatment came to a '*natural end*' coming at the '*right time*'. This may have been because they felt better, that they were in a position to be able to cope on their own, felt it was time to move on and be independent, they had '*found their feet*' or they had achieved the goals originally set out or because the pre-requisite number of sessions had been attended and scores on the measures had improved.

The ending was tentatively mutual for some whereby the practitioner informed the person that '*they were fine now*', that they were '*getting better with the questions*', or their '*scores had come right down*' and the person agreed because they felt better. Occasionally, treatment ended because it was not helping. For some, the ending was not planned or clearly managed.

> *They closed my file because I know I missed the last two sessions or something like that. (NF9)*

Ending treatment was sometimes viewed with anxiety, however patients spoke about the importance of coping by themselves.

> *I feared not being able to go see her, but I tried not to look on her as too much of a lifeline. She said she'd taught me all she could and how did I feel managing on my own, and I thought yes, I wasn't 100% but I could try and she still kept the door open … we spaced out the meetings towards the end. (NM6)*

It was reassuring to be able to telephone the service direct if their problems returned.

> *You go away and then something out of your control happens and it can throw you out of kilter again, and rather than have to start the process again by going back to the GP, then to be referred and so on, I do find, you know, I did find it, erm, what's the word, that, reassuring, that if something did crop up I could phone her. (DF20)*

Whilst some felt able to manage on their own, others did not feel ready for this. In these cases, their sessions were either extended for another minimum period, or the patient was referred to another therapist. Most patients found this helpful. However, it was not always clear to the patient that these options were available, and some were left without a clear understanding of, or being part of making a decision about, what was happening.

> *I just wondered if there was any more help other than this. In some ways I didn't know that I would get referred on to something else and I just thought "is this the best they can offer?"..... he said he would refer me on to maybe a Clinical Psychologist at the end of his treatment and it is all quite a vague memory... When you have been quite depressed you don't quite take it all in as much. (DF13)*

Some stated that they thought they were doing well, or had under played their distress in face-to-face meetings, then suffered between sessions and after

---

termination, resulting in therapy being ended when in fact there had been little improvement.

4.1.31 Reflections on Personal Change

Looking back on their treatment, respondents most commonly reported changes in thinking style. Having a greater understanding of the causes of their problems was important for participants. For people suffering from anxiety attacks the reassurance that their symptoms were not related to a physical condition, and for those suffering with pain having an insight into the interplay between physical and psychological symptoms was important.

For a few, insight into how their behaviour affected their mental state was important, and also aided them in the recognition of triggers for their emotions.

> *I sort of identified that a lot of my anger came from anxiety and episodes of depression ... it's a really interesting process to sort of unpick things that I've done for many, many years and probably become a sort of set behaviour and I wasn't even necessarily aware of. (NF2)*

For others, the reassurance that what they were feeling was normal under the circumstances, that they were not '*going crazy*', that they were not '*a freak*', that other people had similar problems, and they were not on their own, was valued. This was something they realised either by talking to the practitioner, or reading the booklets or doing BtB.

Many respondents simply stated that the treatment/practitioner helped them to view things differently, to see things '*in a different light*'. No longer having feelings of guilt and thinking that they were to blame was an important change expressed by some. This could be either specific to a particular event or more generally that things going wrong were not always their fault, or that they were not responsible for other people's behaviour.

> *[Therapy] helped me understand and accept my childhood, it wasn't all negative, there's a reason why things happen. (NF8)*

For some therapy was also a means of self-discovery; change included feeling less of '*a victim*', or '*a failure*' and having greater confidence.

> *I got a lot better and a lot more confident, I'm a different person… I found out I could do me job…. because I really panicked [when] I thought I wouldn't be able to work, wouldn't be able to do it again. But I could, and I*

---

realized the job was never the problem… It was all the other mess going on. (DF11)

As a result of a change in thinking style many felt that they had become more positive, both as a person and in their outlook on life, not to always look '*on the negative side*' '*jumping to conclusions*' or that they are being judged in a negative light by others. The most commonly reported symptomatic change was an improvement in mood: '*don't feel as down*'; '*feel a new lightness in me*'. Respondents reported feeling less angry, irritable or fidgety and instead feeling more relaxed.

I wasn't as angry with everybody and little things would just .., anything that made me snap before, I sort of started to sort of let go of things.(NF10)

Finally, being able to sleep was an important change mentioned by quite a few participants. A high proportion of those who mentioned an improvement in their sleep patterns had had problems at work but now stated that they were '*able to shut work out completely*' or no longer had '*things on their mind*'.

4.1.32   Work Status

In this sample the biggest work changes since IAPT referral were an increase in numbers working full-time and a decrease in numbers on statutory sick pay. There also appeared to be a small drop in the overall numbers in the sample claiming out of work benefits, or not working at all (see Table 11 below).

Table 11. Employment Status of Sample*

	On referral N/77	At time of interview N/77
Employed, but on SSP	12	7
Working Full Time (inc phased returns)	10	20
Working Part Time	4	6
On benefit (IS or IB)	17	14
Unemployed (JSA)	6	4
Maternity leave	-	1
Student	2	3
Retired	1	3
Not on benefits, not working	7	5
Missing data	12	15

*Constructed from service data and interview data

A number of individuals were in the process of, or had recently completed, phased returns to work. Some were pleased.

> *[The GP] just gave us one more sicknote for two weeks, so obviously I was straight on the phone to work, 'yes, I can come back!' It was good. … So I was over the moon with that. …. It's not just the money, like, but it helps. (DM6)*

For others, the return to work was hard.

> *So when I went back to work that was very traumatic – my first day at work they did a return to work interview – I had walked in and within 10 minutes they had me in an office doing a return to work interview and at that point – erm – when I came out I was – it was horrendous. I mean they were very nice and very supportive but I just wasn't ready for it ….I was actually in tears and, erm, I said 'I don't want to be here'. (DF6)*

Sometimes return involved discussion and re-negotiation around former roles and responsibilities.

I then went and had a meeting with the Principal … because our department was closing down and I more or less went to tell him the reasons why I was leaving and why I was unhappy and erm – but at that point he sort of said 'well, you know what about doing this particular - and offered me a different post which I felt more happier about. (DF6)

A number kept going in their jobs throughout the period of IAPT treatment, particularly if they were using BtB. Another group in the sample were those currently not working who had jobs to return to, or who were actively planning such returns. Others were not currently employed, but were quite clear that they wanted to move in this direction, either directly or via a period of further study, or voluntary work.

Some participants associated having to take time off work due to anxiety or depression with a sense of losing control, and feeling guilty or ashamed, particularly where a period of illness had been experienced more than once,

I think I felt guilty that I'd been off for so long. I know in hindsight perhaps I was off for quite a short amount of time really considering I'd quite a serious illness, erm I think it's just the feeling of letting people down, I shouldn't really be off work for this long, I have to get back. (DM7)

The experience of not working also entailed some major struggles for respondents, including navigating the benefit system, dealing with set backs and learning to cope financially.

That has been a big upheaval for me, going from a decent wage to Incapacity Benefit is quite a shock… on the DLA application, they asked about my mental health and I put that I was seeking help, and it does impact on the little things that you have to do everyday, sometimes you cant be bothered to fill in forms, you can't face doing certain things, certain days, so I put all that on the form.. (DM1)

For some it also meant coming to terms with their own limitations, in ways in which they hadn't expected. For example, realising the need to take things slowly and to think about what job might be suitable in terms of their condition. For those able to work, managing some condition-specific symptoms could also be an issue.

I'm always asking questions, stupid questions because I haven't got the confidence in my own decisions anymore and the job I do, I have to make decisions sometimes. I keep making mistakes, I keep having like memory lapses where they've changed things while I've been away and I keep

forgetting they've changed them and go back to what it was like before. (DM8)

Participants indicated the importance of their decision to return to or seek work. It often seemed to involve reassessing former roles and responsibilities and taking steps to address known difficulties, usually through a process of negotiation with employers. The extent to which IAPT workers were involved and/ or supported the process varied.

I went back too early to work, I didn't go back straight away full time. [The PWP] suggested I go on a self-esteem course, held during the day. But there weren't any places until after I was back at work, so that makes things difficult again for people who are trying to work but still have problems. (DF21)

For some the break from work provided a positive opportunity to re-evaluate their work life balance, and to consider not returning to work in the future. For others, returning to work enabled the realisation that they needed a change of direction, or to bide their time until such a change could be possible.

When I went back and realised I didn't actually want to do it, but it wasn't just that I didn't want to do that job, I didn't actually want to work there at all... I ended up training as a chef and I absolutely love it (DF22)

Another group were not in paid employment, and remained so throughout. For some this was sometimes due to long standing health conditions, such as severe epilepsy, because they were full-time students, or because they were carers. Lastly, there was also a small group who chose to leave jobs where, for example, they were bullied, placing their work ambitions on hold and awaiting help from other services.

Fifteen of the sample talked about additional work related help they had received alongside their IAPT experience. A number had seen someone from the employment team and for many this help appeared limited and that they were not seen as a priority.

Referred about 9 months ago and still not getting anywhere. The employment team, forget it. If you are referred and they can't even be bothered, that's annoying (NM3)

Whilst working with the therapist I saw employment service person 3-4 times, but they cannot help me because they can only help certain types of people. (NF11)

Other respondents talked positively about the support they experienced on Conditioned Management programmes and work counselling groups

> *From the Job Seekers and they said would I like to go to the Condition Management Course…, I found that helpful… Well again, it changes the way you think, you know and tries to… like I have had finish work and I am not able to work because of my problems and that was about a nine week session. (DF26)*

4.1.33   Un-negotiated Endings

Of the 77 participants interviewed, 17 left therapy unilaterally (the services recorded 11 of these patients had un-negotiated endings and the remainder identified themselves as having stopped attending the service). Nine of these patients returned final scores showing that they had 'recovered' (using IAPT criteria). Patients indicated a number of reasons for leaving treatment. The majority did so because they felt better, things were improving or they were coping better and so didn't feel as though they needed to continue with treatment any more. However, for many ending treatment early was something they regretted, either because they recognised that they still needed support, or because their problems had returned

> *I thought I didn't need it anymore so I stopped … wish I'd stuck with it because sleep problems have started to come back. (DM6)*

The remaining left treatment early primarily because they did not think it was appropriate to their needs. Reasons given were that they preferred to '*do it alone*', had problems fitting it around work, or '*didn't want to fill things* in'. A few finished because they felt that counselling was more appropriate to their needs at that particular time, either because they preferred to 'offload' rather than 'do things', or because they saw CBT as a foundation from which to seek further help with different issues.

There were instances where participants stated that they *felt 'bad', 'embarrassed' or 'rude'* about finishing treatment early, either because it was done abruptly without contacting the PWP, or because they ignored phone calls. In some cases, they felt that this precluded them from contacting the service again.

> *I don't like wasting people's time and I did leave a message apologising, but I now know that I am not going to get better alone and that I need to try and find help. (DM1)*

Many treatment episodes came to an end because of missed appointments, mix-ups, and miscommunication. Sometimes patients were told, unexpectedly, that their 'case was closed' which resulted in feelings of frustration and confusion, particularly when they could find no record of the practitioner having tried to get in touch. The practitioner leaving the service without the person being informed was also a cause of dissatisfaction.

> I only had the one [meeting] then I got this letter telling me that she had tried several times to contact me…., when I got this letter I checked back to the date…. She [said she] would be in touch with me in four or five weeks time….I checked my phone on the log, ……and it actually went back to about two weeks after we spoke and there was no numbers at all from them, no messages at all and no letters, so it is them fobbing me off. (DM4)

4.1.34   Extent of Therapeutic Change

Those interviewed indicated a range of outcomes with 40 patients returning scores that indicated 'recovery' as defined by the services, 24 had not reached 'recovery' scores, and for 7 their scores had deteriorated (6 did not provide end of treatment scores). About a third of the patients who had not reached 'recovery' indicated that they experienced some improvement. This was either in one area, such as sleeping or depressed thoughts or, they had improved, but some of their difficulties had returned since finishing treatment. A few patients had significant mental health problems and were referred on to secondary services.

Those who indicated they had made substantial gains made comments such as: '*it's been phenomenal,*' '*IAPT has changed my outlook on life,*' '*They saved my life,*' '*I've gone from strength to strength*' and '*I've come a long way considering where I was*.'

Many participants reported some improvement, but did not feel that they had fully recovered or had got back to their former selves and/or still needed support.

> Better than I was when I first started going because … I weren't sleeping and things like that and I'm sleeping a lot better at night. I'm not always in tears - I don't feel as down as I was because I'm getting more sleep. (DF1)

Importantly, for a few, the improvement they had made meant that they no longer felt suicidal and felt more hopeful about the future. Being able to cope better and problems being more manageable was the sentiment most commonly expressed amongst this group.

I couldn't cope with anything - even little things - now I'm getting on top of things. (DF23)

A specific coping skill mentioned by interviewees that they found particularly useful was an ability to recognize the signs that indicated they were starting to go downhill again which meant that they were able to '*self monitor*' their behaviour and '*stop the cycle before it goes on and on'*.

Some participants however indicated that they had made little or no change. In some cases they said they were resigned to having mental health problems for the rest of their lives, or that they were continuing to avoid stressful situations and that their '*difficulties are part and parcel of life'*. Some felt that this type of service simply wasn't suitable for them

Not really sure if there have been any [changes]. I think they're very subtle you know, changes that have happened, you know just thinking a bit more about things (DF2)

I don't think it helped me as much as it could have done; it's perhaps designed for somebody who wasn't in as bad shape as I was. (DF24)

Almost a third of those who had not reached 'recovery' at the end of their treatment had been stepped up or referred straight for CBT. Most found this helped in some areas. However, many of the patients who had been stepped up and had not reached 'recovery' at the end of both treatments said they did not find Low Intensity treatment helpful. For some both the Low or High Intensity treatment was not long enough, or they felt they had waited too long and they needed more than brief interventions. Some had tried BtB and disliked it; others found telephone calls stressful and wanted more personal contact and better communication with their worker. Some said that they needed more than 'advice' and found it difficult to make use of materials they were sent or to engage with the workers' suggestions.

In fact, it might have made me worse in a way because I used to dread her ringing, to point the obvious out- that you should be doing this and you should be doing that. I didn't find it useful.... There was absolutely no chance that I could have ever read anything ever... Because of depression and my memory, words turning into each other. Also I was embarrassed and I didn't really want proof. I didn't want my family coming round and finding notes about anxiety and depression and they weren't aware that I was getting treatment, so I didn't want any reminders of it in the house that I was this crazy lady. (DF7)

4.1.35 Stepping Up

The sense made of stepping up seemed to be linked to a comparison between the experience of Low and High Intensity interventions and practitioners. Patients described diverse experiences of the two levels of interventions. In particular patients spoke about the fit of the intervention to their needs, the relationship with the practitioner, the amount of change they experienced and where they attributed this change. Also patients' sense of power, in terms of desperation, choice and knowledge and whether they had input into the decision to step up influenced their experiences of stepping up.

In general, patients stated that stepping up was largely positive as the Low Intensity Intervention had not resulted in much change, and they attributed this to the intervention not fitting with their needs, i.e. not collaborative or flexible enough and/or to the relationship with their worker.

I felt like I was drowning and someone was throwing a ring to me that landed about ten miles away. (DF13)

I kept thinking that this was not for me and [the PWP] was following what she thought was her thing of treatment but she was forcing me to do these things that I wasn't ready to do (DF)

These people in the main, went onto describe a more positive experience of CBT both in terms of the fit of the intervention and the relationship with their worker. Several participants described their experience of feeling understood and cared for.

[CBT therapist] got an empathy of what I'm experiencing and is able to translate that…she knows I'm trying to do my best. (DF)

Individuals who had not been progressing with the Low Intensity Intervention and had been unaware of anything other than this step seemed to experience hope at being offered CBT.

At least there is some kind of progression to it and the biggest thing for me is if something hasn't worked or not fully worked they have referred me onto something different straight away and that has made a huge difference. (DF13)

Others felt relieved and validated that they were able to step up but also expressed frustration that they had not been offered this sooner and had not been aware of the service structure earlier.

I wasn't aware…that there was an opportunity to have face-to-face contact. If I had known I would have asked for that in the first few weeks. (DF1)

Where people had a positive relationship with their PWP, they expressed some anxiety about being passed on from someone with whom they had a good relationship.

I thought if I can't talk to him like I talked to [the PWP] then I thought I am going to end up back at square one again. That was the scary part. (DF)

A few participants described stepping up as part of a planned package of care and they had positive experiences of both steps.

It was fine for me because it was a different issue... in this particular instance I was starting again because it was a different therapist for a different illness. (NF1)

Many who had not experienced the Low intensity interventions as helpful felt there was some hope for them when they were offered CBT and felt that they were being 'listened to' that the self-help was not working. These experiences related to patients' sense of powerlessness, which came from their desperation for help, and their experience of a lack of choice and knowledge.

I'm trying to make this work so am I allowed to say 'no I'm sorry this isn't working for me, I don't want it anymore?' Or have I got to wait for [practitioner] to say 'I don't think this is right for you and I'd like to pass you on to someone else'. (DF)

4.1.36   Limitations of this analysis

More patients were interviewed in Doncaster than in Newham, which may have produced some biases in our findings. This happened partly because two postgraduate student studies were conducted at Doncaster only. Our intention was to conduct a small study of the experiences of patients who had received therapy in a language other than English from the Newham service. Although great efforts were made we only recruited one person and were, therefore, unable to consider the views of service users who had accessed this service.

In addition our sample opted-in to the study and although we ensured that the sample included people from a range of ages, outcomes, initial symptom severity etc., we may not have captured the full range of patients accessing the two services.

In terms of quality control we did not seek respondent validation, and analyses by a service user researcher was completed on only some charts. However, the primary analysts discussed all chart summaries and analyses, including our own roles and potential biases.

4.1.37　Summary

There are a number of themes that run through the preceding analyses, which are briefly summarized here as they relate to the study's aims.

4.1.38　Patients' experiences at different stages of the services

- For many patients the first contact with their GP and with the IAPT service was important in helping to identify the problem, provide hope and a way forward. However, it is interesting that many patients did not seek help until they were at crisis point.
- Choice was a central issue throughout the patients' journeys. The first contact with GPs and IAPT services was enhanced when patients had a sense of control and choice and were seen quickly. However, some patients experienced little or no choice in either referral or treatment options.
- Self-referral often gave patients greater self-confidence and hope.
- Information (verbal and written) that could have helped in decision-making was often not available to patients.
- Good communication at all levels helped patients engage with the service.
- Patients' experiences of Low Intensity work varied. This appeared to be for a number of reasons, such as current levels of patient distress, the choices of interventions available to services, the clarity and manner in which the choices are offered, the skills and responsiveness of the PWP.
- Patients indicated that they valued working with PWPs in a structured format. This was particularly so if the interventions were tailored to patients' needs. However, sometimes patients indicated that the self-help booklets were not pitched at the right level.
- Some patients liked the freedom of telephone contacts and the computerized self-management packages, but many found them problematic. Careful introduction, some one-to-one sessions and personal support helped improve the value of telephone/computer working.
- High Intensity work was valued but often thought to be too short.
- The relationship between patient and practitioner was better if the practitioner was responsive, flexible, and respectful. However,

sometimes practitioners were experienced as impersonal or lacking in some way
- A lack of continuity (due to staff turnover) or follow-up was problematic for patients.

4.1.39   Patients' experiences of changes in work status
- Following treatment more of the sample were in work than at the start of treatment and fewer were on benefit
- Routes back to work included phased returns, renegotiating new responsibilities, changing roles with the same employer and changing jobs to a different employer, increasing and decreasing hours to suit personal circumstances
- Thus, for some, treatment provided an opportunity to reassess their work roles, responsibilities and lifestyle; for others, managing work remained a struggle and source of dissatisfaction
- Patients reported some aspects of work support services, such as those offered under the Condition Management Programme, as being particularly helpful in getting back to work, but for others contact by the service was minimal

4.1.40   The experiences of patients who did not complete treatment
- Patients who did not complete treatment were less satisfied with the service than those who completed treatment
- Some who stopped attending the service regretted this later and found it difficult to ask to be referred again
- A number of endings occurred through poor communication between patients and the service

4.1.41   The experiences of patients who had good and those who had poor outcomes
- Many patients indicated that the service helped in some ways
- For some this was life-changing
- For others, the changes experienced were minimal and these patients felt that they needed help for much longer and more than 'just advice'.

4.1.42 Patients' experiences of stepping up

- For many 'stepping up' was seen as positive, but this was dependent on whether there was a fit between the patient and the intervention, the quality of the relationship with the practitioners, and the extent to which the patient had knowledge of what options there were and experienced choice or felt part of the decision making process.

5 Locating patient outcomes from Doncaster & Newham IAPT services within the broader UK practice-based and trials data/literature

Overview

In this chapter we report patient outcomes from the electronic service downloads from Newham and Doncaster demonstration sites and use benchmarking methods to locate them in the context of other sources of data derived from UK primary care. The comparative sources comprise: (1) a UK primary care dataset focusing on the delivery of counselling and psychological therapies collected prior to the implementation of IAPT; (2) published literature reporting on outcomes from routine UK primary care service delivery of psychological therapies, and (3) published literature reporting on patient outcomes from UK trials that comprise patients representative of those who would now be considered as eligible for IAPT services. Our aim is to provide a broader context within which to evaluate patient outcomes from the two demonstration sites. We focus exclusively on patient outcomes as measured by the PHQ-9 [20] and GAD-7 [21].

Background

There has been a steady increase in the use of benchmarking methods as a means of evaluating the impact of psychological interventions both in the US [22,23] and UK [24,25]. This is reflected in procedures for estimating benchmarks with the purpose of comparing the effectiveness of psychological interventions as delivered in routine settings with their efficacy as established in trials [26]. Manualised treatments are initially tested using a trials design and then the transportability benefits are tested via effectiveness studies in which certain parameters are relaxed (e.g., settings) but not others (e.g., diagnostic group, treatment approach). Whilst effectiveness studies retain certain research restrictions (i.e., the extent to which research requirements drive the selection and treatment of patients), practice-based studies invoke fewer or no restrictions. Benchmarks can be used in a relative context in which a given service can be benchmarked against a pool of similar services in, for example, primary [9] and/or secondary care [27]. They can also be used in an absolute sense (i.e., a standard) as in the Department of Health setting a

target of a 50% patient recovery rate for patient outcomes within the IAPT demonstration sites [28].

In light of the above, we wished to utilise a variety of data sources and information with the aim of populating a range of benchmarks in order to locate the outcomes of the Doncaster and Newham Demonstration Sites in the broader context of practice-based and trials outcomes. Specifically, we asked three questions:
   (1)   Do the patient recovery rates reach the target of 50% as set by the Department of Health?
   (2)   How do patient outcomes from Doncaster and Newham demonstration sites compare with those obtained by pre-IAPT primary care services in the UK?
   (3)   How do patient outcomes from Doncaster and Newham demonstration sites compare with published reports of patient outcomes in UK practice-based and trials-based studies?

5.1.1 Data sources

Question 1 related to a performance target and hence required no additional material. However, in service of Questions 2 & 3, we sought corroboration from the benchmarking literature as to the feasibility of setting a 50% target. The psychiatric literature often uses a 50% reduction rate in patients' intake scores as a target and this was one index of change adopted in a report on the initial 12 months of IAPT service in Doncaster [29]. However, this target refers to the severity of the presenting phenomenon whereas the Department of Health target of 50% relates to the percentage of the population presenting for treatment.

Question 2 required data on UK primary care psychological therapies performance prior to the establishment of IAPT services. We used a pre-IAPT (i.e., non-contemporaneous) historical data set derived comprising 24 NHS primary care counselling/psychological therapy services. This data set utilised the CORE-OM [27,29-32]and CORE System [33]. We refer to this data set at the CORE Practice-based Evidence-2008 (CORE-PBE-2008) due to its being collated in 2008 but spanning psychological therapies as delivered between January 1999 and October 2008. The purpose of this data set was to provide a benchmark for primary care counselling and psychological therapies prior to the IAPT initiative. Accordingly, we did not differentiate between models of psychological therapy given the focus related to primary care services as previously delivered rather than a focus on any specific individual brand of psychological therapy. In addition, we set specific criteria for including services, namely bona fide NHS services, and patients as a function of (a) the

severity of their presenting depression and/anxiety, and (b) duration of presenting problem (i.e., >6 months). In the absence of any formal diagnosis, applying these criteria provided the best probability of similar case-mix as the IAPT samples. Hence, the selected sample of patients would most likely have been candidates for receiving an IAPT service and many of the services did subsequently become IAPT services.

Question 3 required selecting published articles from studies carried out in the UK representing complementary designs. First, we also sought evidence from practice-based outcome studies with particular reference to UK primary care settings focusing on common mental health problems – that is, depression and anxiety. We considered evidence from studies published from 1995 onwards in order to provide closer approximation to treatment interventions incorporated within the IAPT portfolio. Second, we sought reported pre-post outcomes from CBT/CT arms of published trials focusing on UK primary care and care at the primary-secondary interface. In particular, we targeted studies that mapped onto IAPT-type interventions – namely, computerised cognitive behaviour therapy (cCBT), guided self-help (GSH), and brief CT/CBT interventions but also included studies that sampled patients from traditional out-patient (i.e., secondary care) services. We did this being mindful that the IAPT demonstration sites were likely to capture patients beyond the confines of primary care.

In seeking these data and published sources, we took account of several factors. First, given that the IAPT initiative is specific to the UK and, more specifically, to England & Wales, we focused our efforts on deriving comparative data and literature from the UK only rather than internationally. Our reasoning was to maximize the relevance of the comparative evidence so as to minimize the effects of non-UK service selection and delivery.

Second, the focus on patient outcomes relies on the credibility of measures and measurement. The IAPT initiative defined a set of outcome measures [34] and the logical comparators would be other studies using the same measures. However, the condition-specific measures of PHQ-9 and GAD-7 had relatively limited use within the UK prior to their adoption in the IAPT Toolkit. This held for both UK-based efficacy studies and practice-based studies where the leading candidate measures were the Beck Depression Inventory (BDI) [35] or Beck Anxiety Inventory [31] and the CORE-OM [27,30-32] respectively. However, although both the BDI and BAI are condition-specific measures, the CORE-OM is a generic measure of psychological distress and we therefore carried out a check to determine the appropriateness of using comparisons in recovery rates by analysing a selected data set within one of the IAPT demonstration sites whereby patients had completed all three measures.

Indices of the extent of psychological improvement (recovery)

We adopted two key indices of psychological improvement: (i) percentage patient recovery rates for services, and (ii) uncontrolled (i.e., pre-post intervention) effect sizes.

The IAPT documentation adopts the term recovery and defines recovery rates for IAPT services as "*patients moving to below caseness on clinical outcomes scores as a proportion of the number of people ending contact with services and receiving at least two sessions of treatment." (p.26)* [28]. We have followed this definition for consistency with IAPT documentation and operationalised it according to published characteristics of the various outcome measures (see Section 5.4.1 below). However, regardless of the measure used, caution should be exercised in assuming that the status of recovery applies to all patients whose scores pass a specific cut-off point. This issue notwithstanding, the DH has set the criterion of a minimum 50% recovery rate for those patients completing treatment and aspiring to a recovery rate within the range of 50-60% of people for an effective service [28]. Recovery rates provide a summary statistic that is directly meaningful to practitioners, service managers, commissioners, and policy makers.

An uncontrolled (i.e., within-group) effect size (ES) quantifies the difference between the pre- and post-intervention scores using the standard deviation as the unit of spread. Uncontrolled ESs represent the difference *within* a single condition (i.e., pre- and post-intervention scores) in contrast to controlled ESs that represent the difference in outcomes *between* two groups' conditions assuming a similar intake score. Hence, uncontrolled ESs will be logically larger than controlled ESs. Effect sizes are used in academic reports and in meta-analyses and provide a common metric across studies. However, they are a statistical construct and often difficult to determine direct clinical meaning and implication. The standard interpretation of Cohen's controlled effect size *d* is 0.8 is large, 0.5 is medium, and 0.2 is small. The body of literature on effect sizes has retained these interpretations for uncontrolled effect sizes but it needs to be seen in the context of a within group effect rather than a difference between interventions. Hence a within group effect size of 0.8 is still a large effect and equivalent to the score of the average treated patient being better than 79% of patients' scores before the intervention.

5.1.2 Recovery index 1: Caseness criterion

We constructed a range of change indices. The end-state criterion related to specifying a cut-off score below which a patient might be deemed to belong to

a non-clinical rather than a clinical population. For any given measure, there may be a single agreed cut-off score or a variety of options, but the lower the cut-off score, the more stringent the criterion. For the PHQ-9, the acknowledged cut-off score of 10 was adopted [20]. For the GAD-7, the measure developers recommended a cut-off score of 10 [21] but a score of 8 has also been used [36]. In the absence of clear evidence as to which cut-off is more appropriate, we report results on each cut-off initially in order to highlight the impact of adopting one or other of these values. We report the number of patients whose post-treatment score is below cut-off as a percentage of patients whose pre-treatment score was above the cut-off. We refer to this as Recovery Index 1 (i.e., percentage of patients below caseness).

5.1.3 Recovery index 2: Reliable and clinically significant improvement criterion

However, all measures carry a component of error and are characterised by differing levels of reliability. The definition of moving below caseness does not include the requirement that the change achieved by any individual patient should exceed the measurement error inherent in any measure. To accommodate this point, we have calculated recovery rates with and without an index of reliable change [37]. The index of reliable change is a function of the reliability of the specific measure and the variance of the sample [38]. Two indices of the reliability of any measure are (i) test-retest reliability, and (ii) internal reliability (i.e., alpha co-efficient). The former is the most appropriate but is also, invariably, lower than the corresponding alpha co-efficient. A lower reliability value results in greater measurement error with its inherent suppression of recovery rates. This is likely to yield a more conservative and cautious index of recovery. The primary purpose of this procedure was to ensure that change due to the inherent unreliability of any measure is not attributed as change in a patient's psychological state and thereby inflating the effectiveness of the intervention.

We followed standard procedures and derived indices based on the test-retest reliability of the measures from the measure developers together with properties of our samples [20,21,38]. For the PHQ-9 these yielded a change score of 7 based on the reported test-retest reliability of .84 [20] and 6 for the GAD-7 based on a test-retest reliability of .83 [21]. We took this criterion of degree of change together with the requirement of passing from being a case to a non-case (as defined above) as our second recovery index (Recovery Index 2).

5.1.4 Uncontrolled effect sizes: Secondary recovery index

The advantages of the effect size (ES) are that it can be applied across measures and can be used to derive a number of statements enabling comparisons with the research literature. The disadvantages are that because the index is neutral, it lacks clinical meaning and appears too abstract for practitioners or service managers. It is also highly sensitive to the value of the SD that is the denominator in the calculation such that any restriction in the sampling range (which results in a smaller SD) thereby yields larger ES. There is also inconsistency as to which SD is used for calculating uncontrolled ES (calculated by deducting the post score from the pre score and dividing by the SD). The pre-score SD is generally recognised as being the most appropriate and conceptually clear and this is the one used in this report. By contrast, pooling pre- and post-therapy SDs within the same group or using the post-therapy SD make gains in relation to the pre-treatment status difficult to ascertain. Although these appear fine details, they can have a considerable impact on the resulting ES. We took the effect size criterion as an additional secondary recovery index.

Data and literature sources for derivation of benchmarks

5.1.5 Comparisons with pre-IAPT UK primary care mental health outcomes

We derived our benchmarks from a UK database drawing on 35 predominantly primary care services comprising a total of 70,245 patients. Of these, 68,239 were aged 16 and over and had been assessed for therapy. From these, 39,324 received one-to-one treatment (2 or more sessions, including assessment) and of these 14,271 were rated as 3-4 on a 4-point scale of Anxiety Severity or Depression Severity and for a duration of > 6m. Of the 35 services, 32 sites had return rates >50% and of these 24 were NHS sites. This yielded a patient sample of 7,939 patients of which 5,709 completed their treatment. We labelled these as Benchmark Sample 1 and 2 respectively. The intake mean CORE-OM score for all patients (Sample 1) was 20.2 (SD=6.0) and for those completing treatment (Sample 2) was 19.8 (SD=6.0). These CORE-OM scores are slightly higher than those reported from Doncaster (M=18.8; SD=5.9) and Newham (M=18.3; SD=6.1) in an earlier account of the demonstration sites [36]. Hence, applying criteria of severity and duration yielded a sample that appeared to be broadly comparable with those patients

presenting to the two demonstration sites, at least as determined by our benchmark measure.

In selecting a return rate of >50%, we investigated a range of return rates mindful of the IAPT requirement of a 90% return rate for services and the assertion that lower return rates yield an inflated recovery rate in the order of 5% per 10% drop in return rate. We calculated uncontrolled effect sizes for a range of return rates at the level of sites as follows: >50% return rate = ES 1.74; >70% return rate = ES 1.72; >85% return rate = ES 1.58. However, at the level of therapists, the effect sizes for a range of return rates were as follows: >50% return rate, ES= 1.75; >70% return rate, ES = 1.78; >90% return rate, ES = 1.94. In light of the small N for the higher-return sites, we selected the sample yielding the largest N of patients (i.e., 50%).

We calculated recovery rates using services as the unit of analysis. We employed a primary cut-off score of 10 based on a report indicating this to be the cutting score between clinical and non-clinical populations and based on an analysis of the Psychiatric Morbidity Follow-up survey [39]. This score is an index of general psychological distress and a score greater than the index would be indicative of a need for psychological support. We reasoned that this was a proxy for the IAPT index that combined information from both depression and anxiety measures (PHQ-9 and GAD-7).

In service of making comparisons between the PHQ-9/GAD-7 and CORE-OM as a benchmark, we tested the agreement level on recovery rates when applying the recovery criterion from the IAPT and benchmark measure configurations. We utilised data from the Newham IAPT download as completion of the data on the CORE-OM yielded a more substantive data set for this purpose. We selected all cases where two or more CORE-OM measures had been completed as comprising the first and last measure. Of these, 402 had a PHQ-9, GAD-7 and CORE-OM completed at the same session (session date and CORE date matched on first and last measure). To be able to determine recovery, we restricted this sample to those patients whose scores were above cut-off on both criteria. This yielded a sample of 318 patients who scored above the cut-off point on the CORE-OM and on the PHQ-9 *or* the GAD-7 (clinical cut off ≥10 for all measures).

Table 12 presents the percentage agreement and disagreement between PHQ-9/GAD-7 and CORE-OM recovery rates. Total agreement was 78.6% (agreements on recovered or not recovered). Importantly, only one patient (0.3%) achieved recovery on the CORE-OM criterion but not the combined PHQ-9/GAD-7 criterion. By contrast, 67 patients (21.1%) achieved recovery using the PHQ-9/GAD-7 criterion but not on the CORE-OM. Hence, recovery rates yielded by the CORE-OM from the practice-based data set used for

deriving benchmarks might reasonably be viewed as capturing virtually all patients who would have achieved recovery using the combined PHQ-9/GAD-7 criterion. We repeated the analysis applying the cut-off score of 8 on the GAD-7 but this raised the total agreement rate by less than 2%. Mindful that arguments could be made for adopting a cut-off score of either 8 or 10 for the GAD-7, we present data using both when reporting recovery rates.

As a further check, we plotted sensitivity against 1-specificity and produced ROC curves that showed above acceptable (>0.70) discriminatory ability for predicting recovery on the PHQ-9/GAD-7 combined criterion from recovery on the CORE-OM (AUC=0.81; 95% CI 0.76 to 0.86) and predicting recovery on CORE-OM from PHQ-9/GAD-7 (AUC=0.83; 95% CI 0.79, 0.88). We also considered parallel benchmarks using the condition-specific cut-off score of 13 as a depression-specific index based on a report comparing the CORE-OM with the PHQ-9 and showing a CORE-OM score of 13 to be the best equivalent to the PHQ-9 score of 10. However, to lessen the complexity we report here only results using the more generic CORE-OM cut-off score of 10.

Table 12. Agreement and disagreement rates for recovery when applying CORE-OM versus PHQ-9 and GAD-7 cut-off scores

			CORE-OM (cut-off=10)		
			Recovered N (%)	Not recovered I (%)	Total
Measures and cut-off scores	PHQ-9 (10) + GAD-7 (10) N=318	Recovered N (%)	113 (35.5)	67 (21.1)	180
		Not recovered N (%)	1 (0.3)	137 (43.1)	138
		Total	114	204	318
	PHQ-9 (10) + GAD-7 (8) N=324	Recovered N (%)	111 (34.3)	58 (17.9)	169
		Not recovered N (%)	6 (1.9)	149 (46.0)	155
		Total	117	207	324

5.1.6 Comparisons with published UK practice-based literature on outcomes of psychological therapies

We utilised data from a systematic review and critical appraisal of practice-based literature [40]. A total of 10 studies were identified [27,41-48] and we used summary data from this article as a further benchmark yielding an overall uncontrolled effect size (ES) as well as a mean for the percentage of clients meeting criteria for reliable and clinically significant improvement from those studies that reported this statistic. This latter statistic adopted a stringent index incorporating both reliable and clinically significant recovery.

5.1.7 Comparisons with published UK trials literature on outcomes of psychological therapies

We carried out a pragmatic review of the literature on UK trials of cognitive/cognitive-behavioural therapy carried out since 1995. Our aim was to capture studies that reflected components of the IAPT portfolio of interventions. Accordingly, the studies reflect a more diverse selection of studies than might normally be gathered within a review. Ideally we sought studies using the PHQ-9 and/or GAD-7 but were mindful that UK studies using these measures were very few. Hence we restricted our review to focus on the BDI-I/BDI-II and/or BAI for depression and anxiety respectively. The search terms were as follows: Depression; Anxiety/or exp Anxiety Disorders/; Stress, Psychological/; common mental health.tw.; Dysthymic Disorder/dysthymia.ti,ab.; depression.ti,ab.; anxiety.ti,ab.; (psychotherapy or psychological therap*).ti,ab.; Counseling/; counseling.ti,ab.; Cognitive Therapy/; (cognitive behavioural therapy or cbt or ccbt).ti,ab.; behaviour therapy.ti,ab.; hypnosis.ti,ab.; psychoanalytic therapy.ti,ab.; psychotherapy/ (see Appendix to Chapter 5 for the full search history).

In addition, we searched the NICE guidelines for Depression [49] and for Anxiety [4] as well as relevant Cochrane Reviews for anxiety disorders [50-52] and the report on Anxiety made to the Comprehensive Spending Review [53] We also considered the Cochrane Review on PTSD [54] that identified 3 UK studies but only one used the BAI [55]. However, there was insufficient commonality of measures to derive a stable overall effect size. In addition, we supplemented this strategy with a hand search of UK journals focusing on primary care mental health of which the main target journal was *Behavioural and Cognitive Psychotherapy*.

A total of 16 studies were identified and we grouped the effects according to the two outcome measures identified with the BDI-I & BDI-II yielding 8 effects from 8 studies [55-62] and the BAI yielding 9 effects from 8 studies [32,62-68][28].

---

[28] One study [58] yielded two effects for anxiety and another [67] yielded effects for both the BDI and BAI.

Calculation of benchmarks

5.1.8 Recovery benchmarks: Practice-based data

Table 13 presents descriptive data for 24 bona fide NHS primary care services with PRRs and their 95% CIs. We reasoned that first-last measures and last observation carried forward (LOCF) analyses were premised on more and less optimistic outcomes respectively. Hence we generated separate benchmarks for these two data approaches. The resulting percentage recovery rates (PRRs) for the first-last measures analyses for Samples 1 and 2 were 58.1% and 58.6% respectively with CIs approximately 43% to 74%. The LOCF analyses yielded rates of 44.0% (Sample 1) and 54.1% (Sample 2). Based on the assumption that these rates represented the range of likely recovery rates, we took a mid-point between the highest and lowest rates for each of the two samples. We constructed respective CIs using the range common within each pairings (i.e., the range that each pair of CIs encompassed). Accordingly, this approach generated the following values:

- *Sample 1*: mid-point (benchmark) = 51.1% (CI = 43.3% to 58.9%)

- *Sample 2*: mid-point (benchmark) = 56.4% (CI = 43.2% to 69.7%)

Table 13. Service benchmarks from NHS CORE practice-based data set (>50% pre-post return rate for service; n=24)

Benchmark sample	Analysis	N (%) pre-CORE-OM	N (%) in clinical range at first contact	Percent recovery rate (PRR) (95% CI)
Sample 1: All patients with 2 or more contacts (n=7935)	First-last measure	5187 (65.4)	4915 (94.8)	58.1 (43.3 to 73.0)
	LOCF	7443 (93.8)	7110 (95.5)	44.0 (29.0 to 58.9)
Sample 2: All patients with an agreed ending (completers) (n=5709)	First-last measure	4953 (91.1)	4693 (94.8)	58.6 (43.2 to 74.1)
	LOCF	5438 (95.3%)	5156 (94.8)	54.1 (38.6 to 69.7)

5.1.9 Recovery benchmarks: UK practice-based review

A systematic review of practice-based studies [40] reported a mean recovery rate for 7 studies sampling primary care patients of 56% (range: 45% to 64%; median = 58%) based on completed cases and not specifically CBT interventions. Importantly, the recovery rate included the additional criterion for a score meeting the reliable change index and is therefore consistent with the more stringent recovery index of reliable and clinically significant improvement. Although predominantly drawn from studies using the CORE-OM, there was no overlap in data sets reported on in the previous paragraph. The recovery rate closely approximated that from the preceding analyses of CORE data.

5.1.10 Uncontrolled effect size benchmarks: UK practice-based data

We used the same systematic review reporting on practice-based studies of face-to-face psychological therapies [40]. A meta-analysis was carried out using STATA ('Metan' command) [69] and the resulting forest plot is presented in Figure 7 in which the output shows, for each study, the treatment effect together with the corresponding 95% confidence interval and the percentage weight contributed to the overall meta-analysis. The meta-analysis yielded an overall uncontrolled ES = 1.29 (95% CI 1.26 to 1.33). This value closely approximated the ES arising from Sample 1 LOCF analysis in the CORE service data (i.e. uncontrolled ES = 1.22; see Table 14).

For the effect sizes, we selected the most conservative ES from our primary sample – Sample 1 – namely 1.22 together with its associated confidence interval (1.18 to 1.26). Because effect sizes were calculated on the basis of N of patients, the CIs were relatively tight, suggesting discrete distributions. However, our selected ES closely matched that from the systematic review [40]. Although the latter was based on analyses of samples of patients completing treatment, it drew on a wide variety of disparate services, a factor accounting for the significant heterogeneity of studies. We reasoned that the similarity of these two effect sizes, each drawn from different samples and based on differing analyses, provided a plausible estimate of effect sizes that might be expected in NHS primary and primary-secondary care services. An ES of 1.22 equates to the 89th percentile.

Table 14. Uncontrolled effect size benchmarks: UK practice-based data set

Benchmark sample	Analysis	N (%) pre-CORE-OM	CORE-OM		Uncontrolled Effect Size (95% CI)
			Pre-treatment score	Post-treatment score	
Sample 1: All patients with 2 or more contacts	First-last measure	5187	19.81 (5.93)	9.46 (6.52)	1.74 (1.70 to 1.80)
	LOCF	7443	20.23 (5.95)	13.02 (8.32)	1.22 (1.18 to 1.26)
Sample 2: All patients with an agreed ending (completers)	First-last measure	4953	19.77 (5.92)	9.30 (6.42)	1.77 (1.72 to 1.83)
	LOCF	5438	19.82 (5.96)	10.29 (7.16)	1.60 (1.55 to 1.65)

Figure 7. Forest plot of 10 UK practice-based studies from Cahill et al. [40]
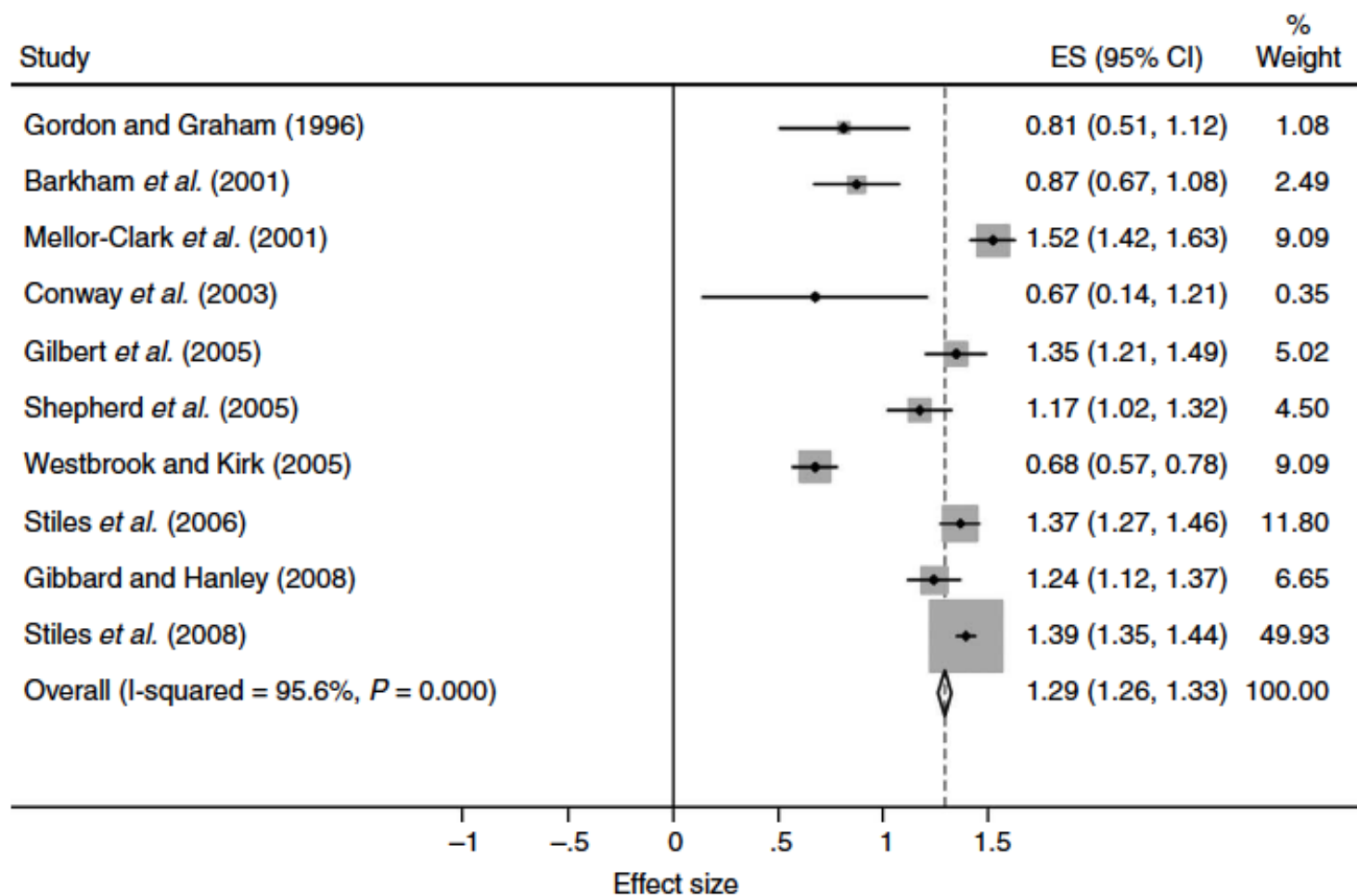
Study	ES (95% CI)	% Weight
Gordon and Graham (1996)	0.81 (0.51, 1.12)	1.08
Barkham et al. (2001)	0.87 (0.67, 1.08)	2.49
Mellor-Clark et al. (2001)	1.52 (1.42, 1.63)	9.09
Conway et al. (2003)	0.67 (0.14, 1.21)	0.35
Gilbert et al. (2005)	1.35 (1.21, 1.49)	5.02
Shepherd et al. (2005)	1.17 (1.02, 1.32)	4.50
Westbrook and Kirk (2005)	0.68 (0.57, 0.78)	9.09
Stiles et al. (2006)	1.37 (1.27, 1.46)	11.80
Gibbard and Hanley (2008)	1.24 (1.12, 1.37)	6.65
Stiles et al. (2008)	1.39 (1.35, 1.44)	49.93
Overall (I-squared = 95.6%, P = 0.000)	1.29 (1.26, 1.33)	100.00

Effect size

Project 08/1610/154

Figure 8. Forest plot of 8 UK trials of CBT-based interventions for depression (BDI)

Study ID		ES (95% CI)	% Weight
Blackburn (1997)-CBT		0.95 (0.36, 1.55)	8.91
Ward et al.(2000)-CBT		1.77 (1.28, 2.27)	12.68
Proudfoot (2003)-cCBT		1.20 (0.73, 1.67)	14.03
Brown et al.(2004)-SH		0.78 (0.31, 1.25)	14.13
Salkovskis et al(2006)-SH		1.43 (0.88, 1.99)	10.14
Lovell et al.(2008)-SH		1.10 (0.37, 1.83)	5.88
Purves et al.(2009)-cCBT		1.72 (1.23, 2.20)	13.53
Kessler et al.(2009)-cCBT		2.20 (1.82, 2.59)	20.70
Overall (I-squared = 76.4%, p = 0.000)		1.49 (1.31, 1.66)	100.00

Effect size

Figure 9. Forest plot of 9 UK trials of CBT-based interventions for anxiety (BAI)
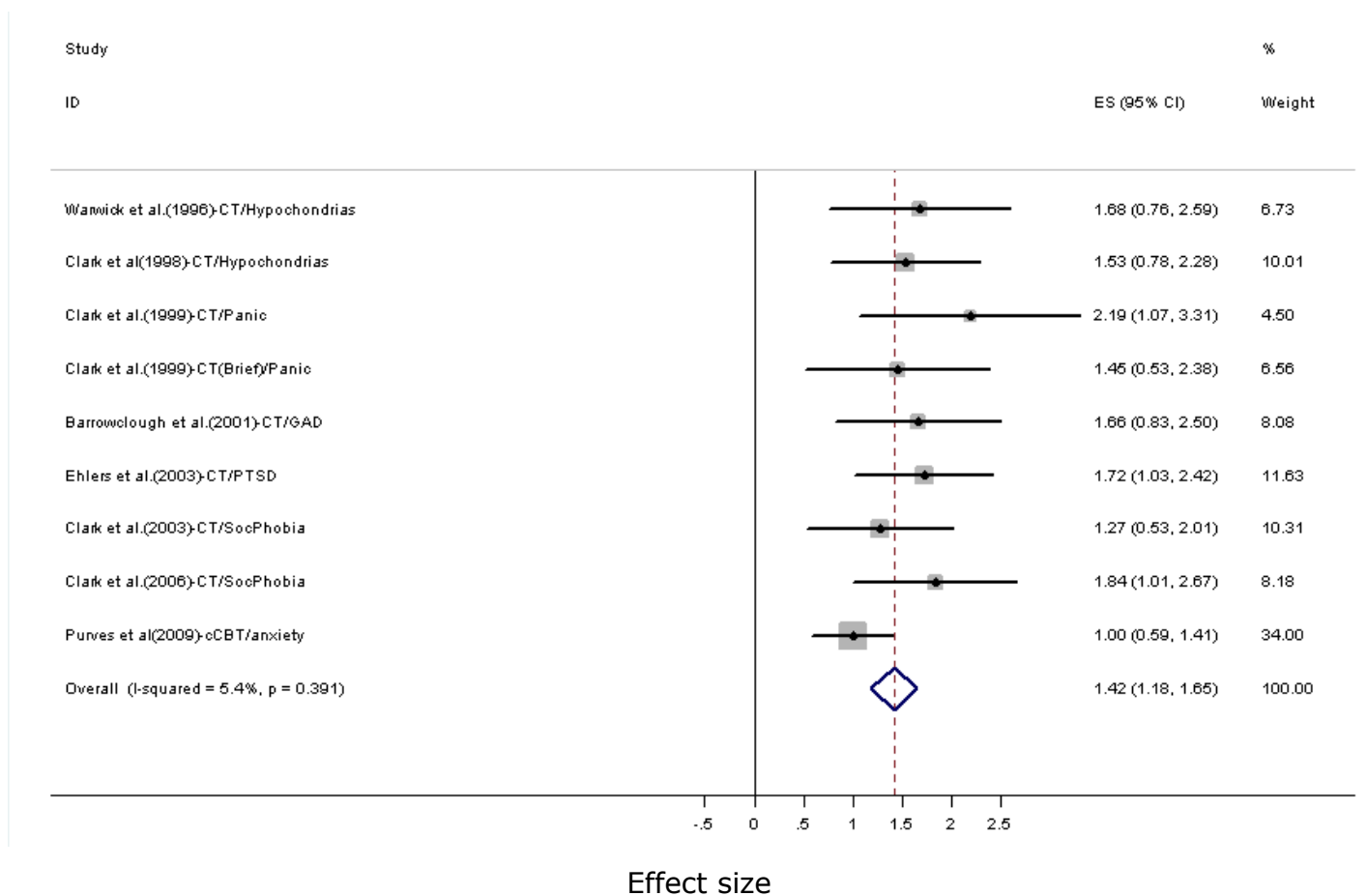
Study ID	ES (95% CI)	% Weight
Warwick et al.(1996)-CT/Hypochondrias	1.68 (0.76, 2.59)	6.73
Clark et al(1998)-CT/Hypochondrias	1.53 (0.78, 2.28)	10.01
Clark et al.(1999)-CT/Panic	2.19 (1.07, 3.31)	4.50
Clark et al.(1999)-CT(Brief)/Panic	1.45 (0.53, 2.38)	6.56
Barrowclough et al.(2001)-CT/GAD	1.66 (0.83, 2.50)	8.08
Ehlers et al.(2003)-CT/PTSD	1.72 (1.03, 2.42)	11.63
Clark et al.(2003)-CT/SocPhobia	1.27 (0.53, 2.01)	10.31
Clark et al.(2006)-CT/SocPhobia	1.84 (1.01, 2.67)	8.18
Purves et al(2009)-cCBT/anxiety	1.00 (0.59, 1.41)	34.00
Overall (I-squared = 5.4%, p = 0.391)	1.42 (1.18, 1.65)	100.00

Effect size

5.1.11   Uncontrolled effect size benchmarks: UK trials literature

The 16 identified UK studies reflect various intervention models representative of psychological interventions that are components within the IAPT services. We carried out a meta-analysis using STATA [69] for the BDI and BAI measures separately and the results are presented in Figures 8 and 9 respectively. For depression, the overall effect size was 1.49 (95% CI 1.31 to 1.66) and for anxiety it was 1.42 (95% CI 1.18 to 1.65). These effect sizes equate to the 93rd and 92nd percentile respectively. The BDI study sample yielded a significant heterogeneity due to the effects of two studies [55,61]. For example, excluding the study by Kessler et al [58] reduced the ES to 1.30 (95% CI 1.10 to 1.50; heterogeneity p=0.04). We present differing groupings of these studies in the Appendix to Chapter 5.

Method

5.1.12   Data sets

The data sets were defined as electronic service downloads as of 1st May 2009 (Doncaster) and 31st March 2009 (Newham) and comprised the total patient sample as previously reported (see Chapter 2: Patient flow diagrams). The patient samples used in the present analyses comprised closed cases yielding a total patient N of 4616 patients from Doncaster and 1422 for Newham. Figure 10 presents the respective Ns in the patient flow for Doncaster and Newham.

5.1.13   Primary and secondary samples

From the total N of closed cases within each service, three distinct subgroups of patients were constructed to match specific research requirements.

Sample 1: This comprised patients who had received at least 2 sessions within the IAPT service of which one session was an assessment and whose case was closed[29]. This requirement made the sample closest to the requirement as set

---

[1]For the Newham sample cases were selected who had at least one intervention session, the rationale for this was that measures were often completed at 'flexible engagement sessions'. Selecting those who had received 2 sessions would include many who had received no intervention. However, this sample does exclude those who had more than one assessment session, but no intervention (which may be included in the Doncaster sample).

out in the IAPT documentation [28]. Accordingly this sample excluded patients who had received an assessment only. The N of patients was 4154 for Doncaster and 1148 for Newham. We identified this as our primary sample, comprising patients assigned to and receiving at least the minimal dosage as set by the IAPT services.

Sample 2: This comprised all patients who completed the agreed intervention within Doncaster (n=2932) and Newham (n=592). This sample therefore comprised a subset of sample 1 and defined as those who adhered to the intervention and received the maximum dosage as appropriate for each patient.

Sample 3: This comprised all patients who were closed cases within Doncaster (n=4616) and Newham (n=1422) and defined as those patients receiving an assessment and being offered treatment. This sample therefore comprised all patients assigned to the intervention. We termed this an intent-to-treat sample.

5.1.14   Analyses

Within each of the three samples we carried out a standard approach of using first and last available measures with the sample size of each data set defined by number of patients completing both PHQ-9 and GAD-7 at intake. Because of the availability of measures taken at each contact, the end point was taken as the last completed administration of each measure, thereby adopting a last observation carried forward (LOCF) procedure. This would likely produce larger recovery rates for Sample 2 (completers of assigned treatment), where the last measure would be at the end of treatment while recovery rates would be smaller for Sample 3 where all those assessed for a treatment were included. Therefore for those patients that 'dropped-out' after their Assessment, their last scores would also be their intake scores. Recovery rates for Sample 1 are likely to lie between Samples 2 and 3, as they include patients who received an intervention but may have dropped-out without completing the allocated course of treatment.

5.1.15   Service-specific definitions in relation to patient flow

Doncaster: Re-referrals were not included in this analysis as it was not possible to identify the first and last measures of each treatment episode. A total of 492 patients were therefore excluded. The first contact was almost always an Assessment, although treatment may also have occurred and the treatment Step and treatment type (s) was recorded. Each planned contact was recorded in terms of 'Type' (mainly phone or face-to-face), purpose (e.g., Assessment,

'therapy session 1-to-1'), treatment (e.g. cCBT, CBT, information), the outcome (e.g., attended, cancelled, stepped-up) and the duration. Due to changes in what data was collected as well as developments in the data collection systems during the time period, a large amount of this data is missing. The type of treatment ending was recorded with those described as 'completed treatment' considered to have had a 'planned ending' while those that began treatment and subsequently 'declined treatment' or were considered unsuitable or 'drop-out' were termed 'unplanned endings'.

Newham: The Newham dataset used three variables to record activity; 'session step'; 'session type'; session 'outcome'. All activity was recorded as a session whether or not contact was made with the patient, including data entry, supervision, and telephone calls. The practitioner recorded the 'Step' in therapy at which the patient was at the time (e.g., Flexible Engagement, Assessment, Step 2/3), then the type of session (e.g., Flexible Engagement, Assessment, Intervention), and finally the outcome of the session (e.g., attended, DNA, cancelled).

There were large inconsistencies in the way data was recorded for instance it was difficult to ascertain how many sessions patients actually attended. The outcome category 'attended' is intended to record patient attendance only. There were however 361 instances of 'discussion with health professional' that were recorded as 'attended' and these have *not* been included in the analysis. An alternative method of determining contact with the patient would have been to use the 'clinical time' recorded. However, in 1444 instances a clinical time was recorded for a session outcome that was recorded as DNA or cancelled and 859 instances where the session was recorded as 'data entry/clinical note', and 263 instances of missing data.

The type of session attended was also difficult to ascertain; in 1942 instances the session step was recorded as 'flexible engagement' but the 'session type' recorded as an assessment session. In 125 cases the session step was recorded as 'assessment' but the session type recorded as 'intervention'. Additionally, the session step was recorded as Step 2 or Step 3 but the session type was recorded as an assessment session (n=442). For the purposes of this analysis, the 'Session Type' variable considered to be the most reliable variable to determine whether an assessment or intervention session had taken place and was therefore used in this analysis. It should be acknowledged that there might be some inaccuracies in these numbers that are unavoidable and that these may affect the outcome data analysis.

Results of standardised clinical outcomes

In this section we report on the clinical outcomes from Doncaster and Newham services. We present results for the two services in tandem in order to facilitate comparisons against the derived benchmarks. However, due to differences in local procedures in collecting and reporting data, as well as differing electronic patient information systems, certain data are not directly comparable. Data tables focus on the reported rates based on patients originally in the clinical population and on the rates derived from applying the clinical cut-off alone (Recovery Index 1) as well as for reliable and clinically significant change (Recovery Index 2). We report on analyses using the cut-off score of 10 and 8 for the GAD-7. In addition, we report the uncontrolled ES for each measure separately derived from each sample.

Figure 10.   Flow diagrams for Doncaster (left) and Newham (right)
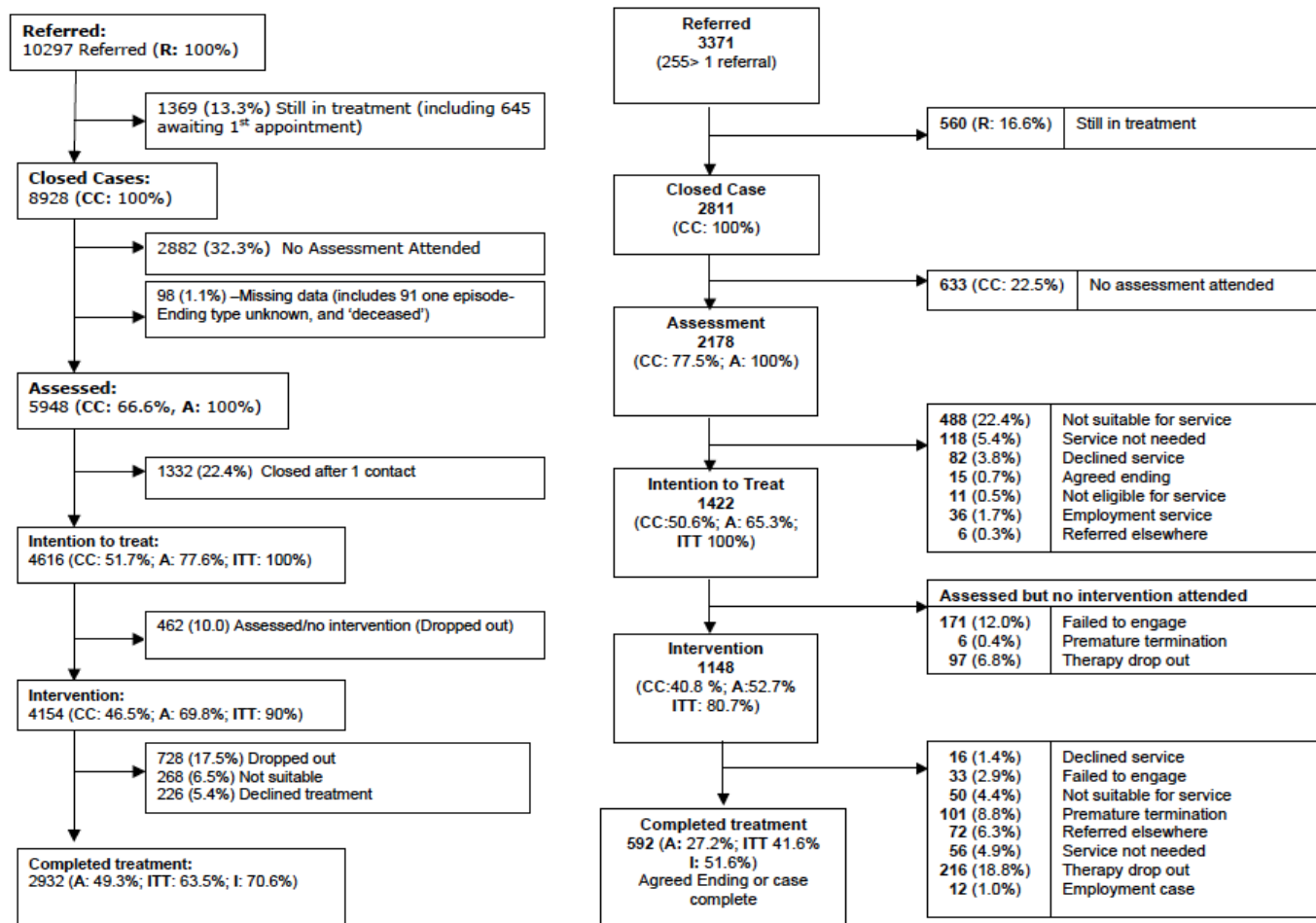
Project 08/1610/154

5.1.16 PHQ-9 and GAD-7 combined recovery rates

Primary sample & analyses: cut-off scores of 10 (PHQ-9) & 10 (GAD-7)

Table 15 presents the percentage recovery rates (PPRs) for the Doncaster and Newham services for analyses using the first and last available measures on the three samples using the combined PHQ-9 and GAD-7 recovery indices. For Sample 1, the percentage recovery rates using Recovery Index 1 (caseness) were 49.4% (Doncaster) and 47.2% (Newham), while the rates for Recovery Index 2 (reliable and clinically significant improvement; RSCI) were 43.7% (Doncaster) and 40.2% (Newham).

Secondary samples and analyses

Table 15 also reports results of secondary samples and analyses. For patients who completed their intervention (Sample 2), the PRRs using Recovery Index 1 (caseness) were 59.4% (Doncaster) and 70.9% (Newham), while for Recovery Index 2 (RCSI) they were 53.9% (Doncaster) and 62.6% (Newham). When all patients assigned to an intervention were analysed (Sample 3), the PRRs for Recovery Index 1 (caseness) were 44.5% (Doncaster) and 40.2% (Newham) while for Recovery Index 2 (RCSI) they were 39.3% (Doncaster) and 34.4% (Newham).

Primary sample & analyses: cut-off scores of 10 (PHQ-9) & 8 (GAD-7)

We carried out similar analyses retaining the cut-off score of 10 for the PHQ-9 but using a cut-off score of 8 on the GAD-7. The results are presented in Table 16. For Sample 1, the percentage recovery rates using Recovery Index 1 (caseness) were 45.8% (Doncaster) and 43.2% (Newham), while the rates for Recovery Index 2 (RCSI) were 40.4% and 36.0% respectively.

Secondary samples and analyses

Table 16 also reports results of secondary samples and analyses. For patients who completed their intervention (Sample 2), the PRRs using Recovery Index 1 (caseness) were 56.1% Doncaster) and 65.9% (Newham) and for Recovery Index 2 (RCSI) were 50.4% and 56.3%. When all patients assigned to an intervention were analysed (Sample 3), the PRRs using Recovery Index 1 (caseness) were 41.3% (Doncaster) and 36.8% (Newham) while for Recovery Index 2 (RCSI) the rates were 36.4% and 30.3%.

Table 15. Percentage recovery rates (PRRs) for Doncaster and Newham service data using first-last measures for PHQ-9 (cut-off 10) and GAD-7 (cut-off 10)

Service & sample	Total N of patients with PHQ-9 & GAD-7 at first contact	N (%) of patients in clinical range on PHQ-9 and/or GAD-7 at first contact	N patients achieving: Recovery index 1[a] Recovery index 2[b]	Percent Recovery Rate (PRR) (95% confidence intervals)	
				Recovery index 1: Below caseness	Recovery index 2: Reliable & clinically significant improvement
Doncaster					
Sample 1 (n=4154): All with 2 or more contacts	4146	3681 (88.8)	1818[a] 1607[b]	49.4 (47.7 to 51.0)	43.7 (42.0 to 45.3)
Sample 2 (n=2932): Completed treatment	2926	2538 (86.7)	1507[a] 1369[b]	59.4 (57.4 to 61.3)	53.9 (52.0 to 55.9)
Sample 3 (n=4616): All accepted for treatment (ITT)	4594	4088 (89.0)	1821[a] 1608[b]	44.5 (43.0 to 46.1)	39.3 (37.8 to 40.9)
Newham					
Sample 1 (n=1148): All with one or more intervention session	1131	972 (84.9)	459[a] 391[b]	47.2 (40.1 to 46.3)	40.2 (37.1 to 43.4)
Sample 2 (n=592): Completed treatment	591	494 (83.6)	350[a] 309[b]	70.9 (66.6 to 74.8)	62.6 (58.1 to 66.8)
Sample 3 (n=1422): All accepted for treatment (ITT)	1338	1153 (86.3)	464[a] 397[b]	40.2 (37.4 to 43.1)	34.4 (29.6 to 35.0)

Table 16. Percentage recovery rates (PRRs) for Doncaster and Newham service data using first-last measures for PHQ-9 (cut-off=10) and GAD-7 (cut-off=8)

Service & sample	Total N of patients with PHQ-9 & GAD-7 at first contact	N (%) of patients in clinical range on PHQ-9 and/or GAD-7 at first contact	N patients achieving: Recovery index 1[a] Recovery index 2[b]	Percent Recovery Rate (PRR) (95% confidence intervals)	
				Recovery index 1: Below caseness	Recovery index 2: Reliable & clinically significant improvement
Doncaster					
Sample 1 (n=4154): All with 2 or more contacts	4146	3767 (90.9)	1726[a] 1521[b]	45.8 (44.2 to 47.4)	40.4 (38.8 to 42.0)
Sample 2 (n=2932): Completed treatment	2926	2605 (89.0)	1461[a] 1313[b]	56.1 (54.2 to 58.0)	50.4 (48.4 to 52.3)
Sample 3 (n=4616): All accepted for treatment (ITT)	4594	4183 (91.1)	1728[a] 1522[b]	41.3 (39.8 to 42.8)	36.4 (34.9 to 37.9)
Newham					
Sample 1 (n=1148): All with 2 or more contacts	1131	1000 (88.4)	432[a] 360[b]	43.2 (40.1 to 46.3)	36.0 (33.0 to 39.1)
Sample 2 (n=592): Completed treatment	591	513 (86.7)	338[a] 289[b]	65.9 (61.6 to 70.0)	56.3 (51.9 to 60.7)
Sample 3 (n=1422): All accepted for treatment (ITT)	1338	1188 (88.8)	437[a] 360[b]	36.8 (34.0 to 39.6)	30.3 (27.7 to 33.0)

Project 08/1610/154

Comparison with benchmarks

We made direct comparisons using Recovery Index 1 based on the primary sample and analysis with the benchmarks established using a similarly derived recovery index from the CORE data. We made these comparisons using the cut-off scores for the GAD-7 of 10 (see Figure 11) and 8 (see Figure 12). Figure 11 shows the recovery rates for both services based on Sample 1 to fall within the range of the confidence intervals for the benchmarks, albeit the lower region, while the rates for those patients completing treatment (Sample 2) lie at the upper region of the confidence interval for the benchmark, with Newham showing particularly good outcomes. Figure 12, where a GAD-7 cut-off of 8 was used, shows this pattern to be replicated although with slightly reduced recovery rates for the IAPT services.

Figure 11.   Comparison of Samples 1 and 2 for Doncaster & Newham with pre-IAPT benchmarks using Recovery Index 1: Cut-offs PHQ-9=10, GAD-7=10
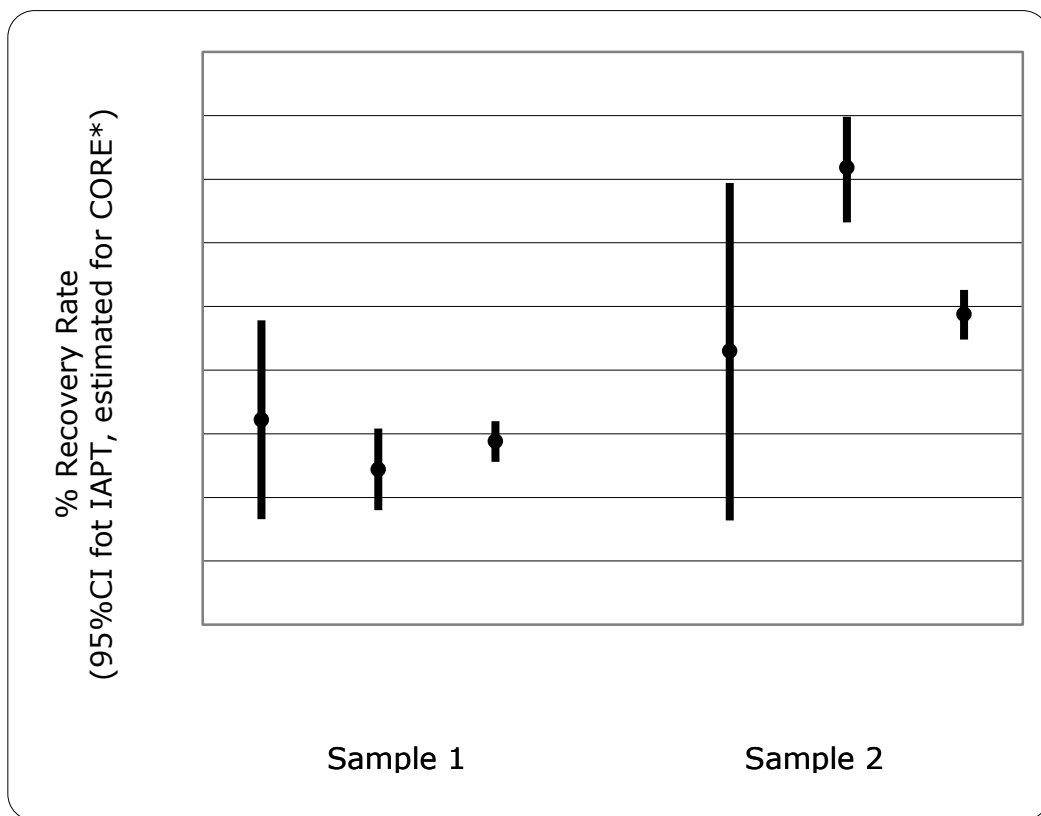
Note: Sample 1 = Patients attending at least 2 sessions (including asssessment);
Sample 2 = Patients completing a defined course of treatment

Figure 12.   Comparison of Samples 1 and 2 for Doncaster & Newham with pre-IAPT benchmarks using Recovery Index 1: Cut-offs PHQ-9=10, GAD-7=8
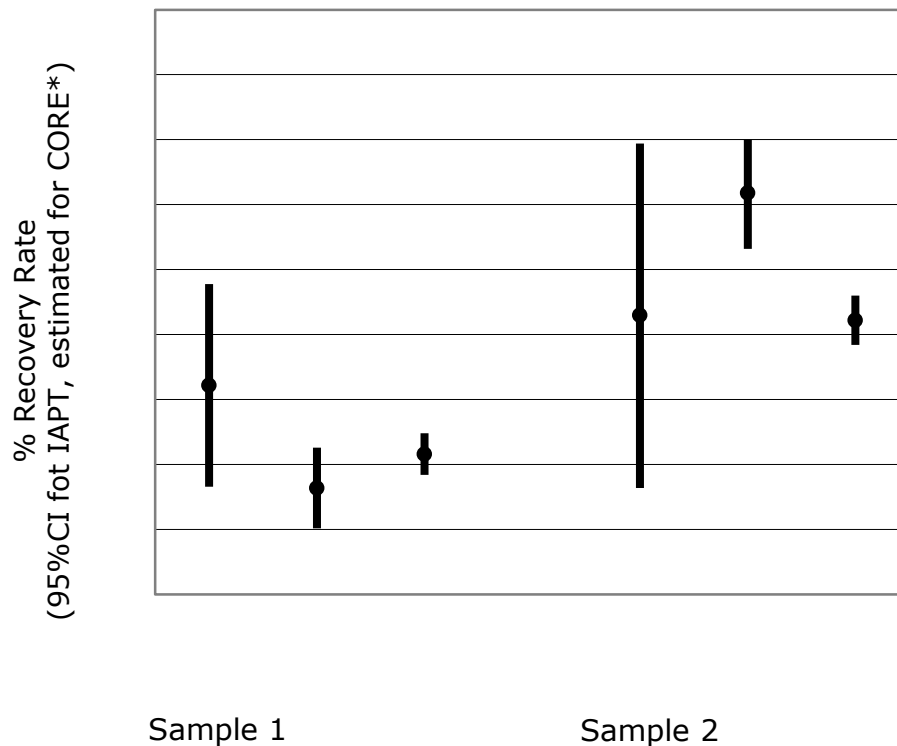
Sample 1   Sample 2

Note: Sample 1 = Patients attending at least 2 sessions (including asssessment);
Sample 2 = Patients completing a defined course of treatment

5.1.17   Phases of rollout in Doncaster and Newham

We tested locally defined phases of implementation to test whether key changes in some part of the implementation affected the recovery rates. In Doncaster, the service managers identified three time periods that signified when the IT data collection systems were developed. The most recent period incorporated the IAPT MDS. The results are presented in Table 17 and show an almost 6% fall in recovery rate from Phase 1 to Phase 2 and Phase 3.

Table 17. Recovery rates for Doncaster across three phases of implementation for patients in the clinical range on initial assessment

Referrals between:	Sample 1: First-last available measure
	Recovery Index 1*
Phase 1: June 06–Nov 06	54.0
Phase 2: Dec 06–Aug 08	48.3
Phase 3: Sept 08–April 09	48.3
Overall	49.4

Note: * Recovery index used cut-off scores: PHQ-9=10, GAD-7=10

Table 18 reports the recovery rates for the three Phases within Newham and shows an 8% drop in recovery rates from 55% at Phase 1 to Phase 2 (47%) and with a further smaller fall to 44% in Phase 3. This drop was evident regardless of the sample or analyses employed. During the first phase the Newham service provided a High Intensity service with experienced (Grade 8) CBT therapists. The second phase was a period of change when lower grade case workers (Grade 5) were employed by the service to provide a Low Intensity service and less experienced CBT (Grade 6/7) therapists were also introduced. During Phase 3 the service was more stable in terms of the relative provision of high and low intensity but experienced a high staff turnover as neighbouring trusts introduced IAPT services.

Table 18. Recovery rates for Newham across three phases of implementation for patients in the clinical range on initial assessment

Referrals between:	Sample 1: First-last available measure
	Recovery Index 1*
Phase 1: Jul 06-Dec 06	55.4
Phase 2: Jan 07-Dec 07	46.9
Phase 3: Jan 08-Mar 09	44.2
Overall	47.2

Note: * Recovery index used cut-off scores: PHQ-9=10, GAD-7=10

5.1.18 PHQ-9 & GAD-7 uncontrolled effect sizes

Uncontrolled effect sizes (ES) for the PHQ-9 and GAD-7 separately are shown in Table 19. For the PHQ-9, patients were included if they had a first measure and a subsequent measure and similarly for GAD-7. On each, those with an intake measure only were excluded.

Table 19. Uncontrolled effect sizes based on PHQ-9 and GAD-7 combined data for samples 1-3 drawn from Doncaster & Newham

Service & sample	Measure	First measure Mean (SD)	Last available measure Mean (SD)	Mean change score	Un-controlled effect size	Per-centile
Doncaster						
Sample 1: All with 2 or more contacts n=4154	PHQ-9	16.08 (6.21)	8.80 (7.26)	7.28	1.17	88
	GAD-7	13.92 (5.14)	7.91 (6.36)	6.01	1.17	88
Sample 2: Completed treatment n=2932	PHQ-9	15.54 (6.33)	7.29 (6.78)	8.25	1.30	90
	GAD-7	13.58 (5.29)	6.60 (5.99)	6.98	1.32	91
Sample 3: All accepted for treatment (ITT) n=4616	PHQ-9	16.13 (6.17)	9.73 (7.53)	6.40	1.04	85
	GAD-7	13.94 (5.14)	8.65 (6.54)	5.29	1.03	85
Newham						
Sample 1: All with one or more intervention session n=1148	PHQ-9	14.65 (6.03)	8.73 (6.98)	5.92	0.98	84
	GAD-7	13.23 (5.13)	7.77 (5.98)	5.46	1.06	86
Sample 2: Completed treatment N=592	PHQ-9	13.85 (6.24)	5.88 (5.67)	7.97	1.28	90
	GAD-7	12.79 (5.32)	5.37 (4.92)	7.42	1.39	92
Sample 3: All accepted for treatment (ITT) n=1422	PHQ-9	14.23 (6.23)	9.96 (7.15)	4.27	0.69	75
	GAD-7	12.66 (5.40)	8.78 (6.16)	3.88	0.72	76

To aid interpretation, the final column provides the equivalent percentile for Samples 1-3 for Doncaster and Newham. For example, the ES of 1.17 for Sample 1 (Doncaster) can be interpreted as the score of the average patient after an intervention being lower (i.e., better) than 88% of patients' scores prior to receiving an intervention. The ESs and their associated percentiles show a consistent ordering of effects with those for Sample 2 (completers) being highest within each site and those for Sample 3 (all assigned to an intervention) being the lowest. Comparisons between the two sites showed the percentiles to be broadly equivalent with the exception of those for Sample 3.

Population sample comparisons

Group comparisons were examined for those patients who met the following criteria: (a) had been discharged; (b) and who were recorded as having one or more intervention session (Newham)/two contact sessions (Doncaster); and (c) completed 2 or more PHQ-9 and GAD-7 measures (Newham n=1057; Doncaster n=4034). ANCOVAs were used to compare PHQ-9 and GAD-7 mean scores, using the first measure score as the covariate. Uncontrolled effect sizes for individual group samples were calculated by dividing the difference between the first and last measure scores by the first measure standard deviation.

5.1.19   Employment status

Table 20 presents the results for employment status and outcomes for both Newham and Doncaster. For both services, patients who were receiving benefits (JSA/IS/IB) had statistically significantly worse outcomes on the GAD-7 than those patients not receiving benefits. For the PHQ-9 effect sizes were smaller only for Newham and one feature of this data is that high PHQ-9 scores at baseline for the Doncaster service.

5.1.20   Self vs. GP referral

Table 21 presents the results for self vs. GP referral for Doncaster and Newham sites. There was no significant difference in PHQ-9 or GAD-7 outcomes for those who did and did not self refer. In Newham outcomes were slightly better for self-referrals while in Doncaster, the reverse was true, although the number of self-referrals in Doncaster was small.

Table 20. Employment status, PHQ-9 and GAD-7 scores, SDs, and uncontrolled effect sizes for Newham & Doncaster sites

	Newham		Doncaster	
Measure	Benefits[1] (N=381)	No Benefits[2] (N=657)	Benefits[3] (N=952)	No Benefits[2] (N=2208)
PHQ-9				
First Mean (SD)	15.95(5.86)	13.87 (6.02)	18.19 (5.42)	15.26 (6.29)
Last Mean (SD)	11.32 (7.27)	7.17 (6.34)	10.96 (7.42)	7.52 (6.82)
Mean Difference	ANCOVA F=64.62 p=<.001		ANCOVA F=62.93 p=<.001	
Effect size	.79	1.11	1.33	1.23
GAD-7				
First Mean (SD)	13.86 (4.93)	12.90 (5.24)	15.11 (4.67)	13.51 (5.19)
Last Mean (SD)	9.76 (6.12)	6.57 (5.61)	9.65 (6.52)	6.90 (6.03)
Mean difference	ANCOVA F=63.82 P<.001		ANCOVA F=78.63 P<.001	
Effect size	.83	1.21	1.17	1.27

[1]Job Seekers' Allowance/Income Support/Incapacity Benefit; [2]Employed/Students/ Retired /Homemakers not on benefits; [3]'Employment benefits'

Table 21. Self and GP referral status, PHQ-9 and GAD-7 scores, SDs, and uncontrolled effect sizes for Newham & Doncaster sites

	Newham		Doncaster	
Measure	Self Referral (N=272)	Not Self-Referral[1] (N=785)	Self Referral (N=52)	Not Self-Referral[2] (N=3981)
PHQ-9				
First Mean (SD)	13.95 (6.07)	14.89 (6.01)	15.77 (6.28)	16.08 (6.21)
Last Mean (SD)	7.99 (6.67)	8.99 (7.07)	9.23 (7.94)	8.79 (7.25)
Mean Difference	ANCOVA F=1.54 P=.26		ANCOVA F=0.46 P=.50	
Effect size	0.98	0.98	1.04	1.17
GAD-7				
First Mean (SD)	12.96 (5.07)	13.33 (5.15)	13.13 (5.50)	13.93 (5.14)
Last Mean (SD)	7.25 (5.96)	7.94 (5.98)	7.65 (6.21)	7.91 (6.36)
Mean difference	ANCOVA F= 1.87 P=.17		ANCOVA F=0.04 P=.84	
Effect size	1.13	1.05	1.00	1.17

[1] GP (726) Other professional (55) Pathways to work (5); [2]GP(3762) Other Professional (172), Social Services (22) Job Centre Plus (25)

5.1.21 Working Age

For Newham those over working age had statistically significantly better outcomes on the GAD-7 but not the PHQ-9 while for Doncaster there was a significant difference on both measures (Table 22), but it should be noted that the numbers of those over 65 was small.

Table 22. Working age and PHQ-9 and GAD-7 scores, SDs, and uncontrolled effect sizes for Newham & Doncaster sites

	Newham		Doncaster	
	Working age (16-65)	Non- working age (>65)	Working age (16-65)	Non-working age (>65)
Measure	(N=1032)	(N=24)	(N=3913)	(N=119)
PHQ-9				
First Mean (SD)	14.74 (6.00)	10.92 (6.32)	16.18 (6.18)	12.82 (6.11)
Last Mean (SD)	8.82 (6.99)	4.96 (5.54)	8.91 (7.28)	5.16 (5.38)
Mean Difference	ANCOVA F=2.43 p=.12		ANCOVA F=10.49 p=.001	
Effect size	0.99	0.94	1.18	1.25
GAD-7				
First Mean (SD)	13.29 (5.10)	10.54 (5.64)	14.00 (5.10)	11.45 (5.90)
Last Mean (SD)	7.86 (5.98)	3.88 (4.75)	7.99 (6.37)	5.16 (5.18)
Mean difference	ANCOVA F=5.91 p=.02		ANCOVA F=7.83 p=.005	
Effect size	1.06	1.18	1.18	1.07

5.1.22 Ethnicity

Table 23 shows no difference in outcomes on either the PHQ-9 or GAD-7 between Black and Minority Ethnic (BME) as compared with White European populations, although the number of BME patients in Doncaster was small.

In Table 24, this analysis is repeated, including White Europeans in the ethnic minorities, in contrast to White British. In Doncaster there was no significant difference between White British patients and other patients (which includes BME and White Other), but the numbers were small for the latter.

Table 23. White and Black Minority Ethnic status and PHQ-9 and GAD-7 scores, SDs, and uncontrolled effect sizes for Newham & Doncaster sites

Measure	Newham		Doncaster	
	BME (N=523)	White European (N=516)	BME (N=13)	White European (N=3817)
PHQ-9				
First Mean (SD)	14.46 (6.03)	14.90 (6.02)	17.77 (5.34)	16.04 (6.22)
Last Mean (SD)	8.57 (7.15)	8.91 (6.85)	11.31 (7.65)	8.74 (7.24)
Mean Difference	ANCOVA F=.12 P=.73		ANCOVA F=0.84 p=.36	
Effect size	0.98	0.99	1.21	1.17
GAD-7				
First Mean (SD)	13.14 (4.96)	13.37 (5.29)	16.15 (4.36)	13.89 (5.15)
Last Mean (SD)	7.55 (6.04)	7.97 (5.93)	10.00 (6.25)	7.86 (6.36)
Mean difference	ANCOVA F=.82 P=.37		ANCOVA F=0.35 p=.55	
Effect size	1.13	1.02	1.41	1.17

Table 24. White British and not White British Ethnicity and PHQ-9 and GAD-7 scores, SDs, and uncontrolled effect sizes for Newham & Doncaster sites

Measure	Newham		Doncaster	
	Not white British (N=624)	White British (N=415)	Not white British (N=28)	White British (N=3802)
PHQ-9				
First Mean (SD)	14.48 (6.02)	14.98 (6.03)	17.04 (5.47)	16.04 (6.22)
Last Mean (SD)	8.50 (7.07)	9.11 (7.0)	10.64 (8.30)	8.74 (7.23)
Mean Difference	F=.86 P=.36		ANCOVA F=1.27 P=.26	
Effect size	0.99	0.84	1.17	1.17
GAD-7				
First Mean (SD)	13.18 (5.01)	13.37 (5.30)	14.29 (5.46)	13.90 (5.15)
Last Mean (SD)	7.60 (6.01)	8.00 (5.95)	9.07 (6.14)	7.86 (6.36)
Mean difference	F=.79 P=.37		ANCOVA F=.85 P=.36	
Effect size	1.11	1.01	0.96	1.17

Locating the patient outcomes from Doncaster and Newham services within the existing evidence-base

We asked three questions using our available data and literature sources: (1) Do the patient recovery rates reach the target of 50% set by the Department of Health?; (2) How do patient outcomes from Doncaster and Newham Demonstration sites compare with those obtained by pre-IAPT primary care services in the UK?; and (3) How do patient outcomes from the two demonstration sites compare with published reports of outcomes in UK practice-based and trials-based studies? We address each question in turn before considering wider issues relating to these data.

5.1.23 Do the patient recovery rates reach the target of 50% set by the Department of Health?

The data show that when only a simple index of caseness is adopted for the PHQ-9 and GAD-7 combined (Recovery Index 1) and using cut-off scores defined by the measure developers, for those receiving a minimum treatment dosage as defined by IAPT (Sample 1), the two services fell only marginally short of the 50% recovery rate (49% and 47% for Doncaster and Newham respectively). However, these rates were reduced by between 6-7% when the more stringent criterion for recovery, as defined in Recovery Index 2 was applied (44% and 40% for Doncaster and Newham respectively). When the cut-off score of 8 was applied for the GAD-7, the respective rates for Recovery Index 1 were in the mid-lower 40 percent range (46% and 43% respectively) and fell by between 5-6% when the more stringent Recovery Index 2 was applied (40% and 38% respectively).

However, for patients who completed their course of treatment (Sample 2), all rates for both sites regardless of cut-off or recovery index, exceeded the 50% target. The least stringent criterion (i.e., cut-off scores of 10 for both measures and Recovery Index 1) yielded rates of 59% and 71% for Doncaster and Newham respectively. When the most stringent criteria were applied (cut-off score of 8 for the GAD-7 and Recovery Index 2), the rates fell to 50% and 59% respectively. The rates for the intent-to-treat sample showed the same pattern but were, logically, lower – in the region of 4-7% - than the rates for Sample 1.

In our view, the more cautious results arising from adoption of the recovery index defined by reliable and clinically significant improvement (i.e., Recovery Index 2) are likely to be a more robust index of patient change. This is because the requirement for a specified amount of change on a measure in addition to achieving a minimum defined end point score is more likely to translate into patients' experiences of psychological gains as opposed to simply passing a defined caseness cut-off score. The widespread adoption of the term *recovery* as the index of patient improvement, with all

its incumbent meaning of improved functioning, necessitates adoption of sufficiently stringent terms when inferring such a state from statistically derived indices. This issue is even more pertinent when the measures used to determine recovery are primarily symptom-based as they lack the broader coverage of patient general and social functioning.

Adopting a revised score for caseness for the GAD-7 (i.e., 8) – a score that has been adopted as custom and practice in the UK – appears a more clinically logical cut-off score as this is approximately mid-point in the moderate range whereas a score of 10 is at the top end of moderate with a score of 11 being in the moderate-severe band. Lack of clarity surrounding the cut-off score reflects a paucity of empirical developmental work using the GAD-7 measure prior to adoption particularly when this index is used as a direct translation to recovery rates.

Noteworthy are the high recovery rates for both services, and Newham in particular (71%), when derived for patients adhering to their treatments (Sample 2). A possible explanation for this is that the Newham service was originally geared to focusing on longer, high intensity interventions. Considerable investment was made in appointing experienced staff - for example taking the extra effort to recruit some staff with a formal psychotherapy qualification in CBT. In this context, the Newham service may have put considerable effort into engaging and retaining patients once they entered the service, in achieving this good result for those completing the intervention.

The reported recovery rates in the present analysis are lower than those previously reported for Doncaster and Newham using the combined measures rate – 56% and 55% respectively [36]. Our results, using the same cut-off scores (i.e., 10 for PHQ-9 and 8 for GAD-7) yielded rates that were 10% less as compared with the earlier publication [36]. This is consistent with the analysis of the three phases within both services that showed an appreciable drop in the percentage recovery rate after the initial phase (albeit using a less stringent recovery index). The drop was more pronounced in Newham where a plausible interpretation would be that the significant changes in the service delivery configuration, beyond that originally envisaged and owned by the service, led to efforts being diffused away from the original vision.

However, the large throughput of patients within the Doncaster site as compared with Newham – approaching threefold for patients receiving a minimum dosage and 5-fold for patients completing treatment – is praiseworthy given the *raison d'être* for the IAPT initiative was that of improving access within the constraint of targeting a 50% recovery rate.

5.1.24 How do patient outcomes from Doncaster and Newham Demonstration sites compare with those obtained by pre-IAPT primary care services in the UK?

An initial yield of the application of the benchmarks, as defined by Recovery Index 1, was endorsement of the 50% recovery rate set by the Department of Health. Our finding of a best estimate of a 51% recovery rate as representing the mid-point between differing analytic approaches places the Department of Health's target in the context of pre-existing IAPT services. Hence it would appear to be a logical evidence-based aspiration.

In terms of the recovery rates based on caseness/not-caseness criterion, patients receiving the minimum dosage fell within the confidence intervals of the benchmarks. This would indicate that, the new services delivered an equivalently effective service even though they were, by definition, new services with interventions being delivered by newly trained practitioners. Notwithstanding the fact that the IAPT demonstration sites received financial support and kudos, this is a considerable achievement within the timeframe.

When a more focused sample was examined comprising those patients deemed to have completed their intervention, the rates for Recovery Index 1 of 59% (Doncaster) and 71% (Newham) both lay at the upper end of the CI range, particularly for the Newham service. Possible explanations for the high recovery rate for people who completed their intervention at Newham are noted later.

Our analyses reported uncontrolled effect sizes on the PHQ-9 and GAD-7 for patients receiving the minimum dosage as follows: 1.17 (PHQ-9) and 1.17 (GAD-7) for Doncaster and 0.92 (PHQ-9) and 0.99 (GAD-7) for Newham. These showed the effects for Doncaster, based on first-last measures, to be broadly consistent with our conservative benchmark while those for Newham were marginally lower. However, the samples at the demonstration sites comprised all patients in receipt of at least a minimum dosage. When patients who completed their agreed intervention were considered, the effect sizes on both measures at both sites exceeded our conservative benchmark.

Scrutiny of the PHQ-9 and GAD-7 first scores showed a consistent pattern with higher intake scores occurring for Doncaster. However, for the patient completer sample, although initial scores were higher for Doncaster, end scores were lower for Newham, a phenomenon that yielded virtually identical effect sizes at each site. That is, while the initial levels differed, the extent of change was very similar. A more comprehensive account of Newham's patient completer outcomes is warranted, as there may be helpful lessons to be learned about how to achieve the overall low scores for completers and their retention in treatment. Alternatively, this pattern of

change may reflect a naturally occurring phenomenon that determines the extent of change available to individual patients.

5.1.25 How do the patient outcomes from the two demonstration sites compare with published reports of outcomes in UK practice-based and trials-based studies?

Our comparison of patient outcomes from the IAPT demonstration sites with UK literature drawn from practiced-based and trials literature places the IAPT services within the effectiveness range of other practice-based evaluations.

In terms of comparisons with the trials literature for depression, the PHQ-9 ESs for both demonstration sites for those patients who completed their treatments (Sample 2) abutted the lower point of the confidence interval for trials (1.31). By comparison, the ESs for the GAD-7 for treatment completers at both sites approached that of the trials literature. For patients who had at least the minimum intervention (Sample 1), only the ES for the GAD-7 at the Doncaster site fell within the CI for the relevant trials literature and none of the ESs for Sample 3 fell within the CIs. The effect sizes in the region of 1.30 for patients completing treatment in the demonstration sites yield a small effect size difference (i.e., <.20) compared to the trials data (1.42 for anxiety; 1.49 for depression). The ordering in which superior outcomes are obtained by trials compared with practice-based studies is consistent with other comparisons made in the literature [70] . However, the extent of the difference appears to be smaller, probably due to the slightly reduced ESs for trials arising from including studies of lower-intensity interventions (i.e., self-help and cCBT).

Caveats and commentary

We identify a number of caveats in relation to the methods adopted and results obtained. All measurement carries error and is an estimate of true change. Our approach was to calculate change using multiple criteria and in doing so we have located the patient outcomes within the context of pre-IAPT services and current UK-based literature focusing on primary and primary-secondary care services.

5.1.26 Calculating change

The analysis of first and last available scores derived from administration at each session yields a high return rate for the last measure. It also provides a better basis for applying LOCF analysis as compared with its application to pre-post scores that, in effect, takes the first score as the last. The effort afforded in requiring patients to complete sessional measures needs to be matched with the application of more advanced statistical analyses using

multilevel modelling that acknowledges that patients are nested within practitioners who are, themselves, nested within services. However, we are mindful of a number of issues that need to be addressed. First, the sessional use of the PHQ-9 and GAD-7 requires the patient to consider the previous 2-week period. If they are completing this on a weekly basis, there is potential for confusion by the patient as well as statistical issues regarding the same 7-day period effectively being rated twice. Second, the statistical effects of autocorrelation need to be addressed as well as memory and fatigue effects from repeated administration on the same measures. Third, the various mix of interventions for patients in the IAPT services probably makes direct comparisons difficult as such comparators tend to represent single modality interventions. However, our approach has consistently been to locate the current IAPT interventions within the extant knowledge base rather than to infer superiority or otherwise to any intervention. Ultimately, the IAPT models are service-level interventions of which the patient-level outcomes are only one component.

Caution should be exercised in reading and interpreting effect sizes as they are vulnerable to a range of influences. Crucially, they are determined by the size of the denominator and researchers have employed differing denominators (i.e., pre-intervention SD, pooled pre- and post-intervention SD, or post-intervention SD). We adopted the pre-intervention SD as the most commonly adopted procedure and the one that has greatest conceptual clarity. Overall, however, in considering results based on effect sizes, it needs to be kept in mind that these are simply point estimates of the population effect and will vary from study to study. We used percentiles to lessen over-interpretation of small differences between effect sizes, particularly at the upper end (e.g., the adjacent 95th and 96th percentiles comprise the effect size range from 1.60 to 1.81).

5.1.27   Recovery

Finally, we offer two key comments on the results. First, in relation to our recovery indices, a major observation is that the actual rates achieved are determined by various factors, in this instance by the sample used, the cut-off score set for specific measures, and the definition of recovery adopted. In all we considered 12 possible combinations (3 samples x 2 cut-off options x 2 recovery indices) and on only one of these was the target virtually achieved. In considering these combinations, we argue for the adoption of the more stringent recovery index – Recovery Index 2 – as has been the case in the broader psychological therapies literature e.g., [70].

The use of the term recovery rather than, for example, improvement requires an equally stringent threshold to ensure the term retains clinical meaning. In this respect, the strategy of requiring specific scores on both PHQ-9 and GAD-7 is consistent with this approach. We would also support the adoption of 8 as the cut-off score for the GAD-7 in contrast to the

measure developers' original suggestions. Although a score of 10 might be used as a marker for further testing, it seems too high to be used as a marker for recovery. However, to the extent that stringency needs to be built into how recovery is defined, we are mindful that our analyses of the PHQ-9 and GAD-9 against the CORE-OM showed 18% of patients to be deemed recovered on the combination of PHQ-9 and GAD-7 who would not have achieved this criterion on the CORE-OM. The latter measure's inclusion of items on functioning and relationships – reasonable indicators of recovery – may account for this difference and may argue for more than symptom-based measures to be used as the basis for determining recovery. Indeed, while there is clear merit in adopting shorter measures for repeated use (e.g., weekly or sessional), the range of items they contain and the logically higher reliable change index resulting from lower reliability leads to such shorter and symptom-based versions not being optimal in determining recovery.

Beyond these conceptual and methodological considerations, the substantive finding from the analyses is that recovery rates – however defined – are appreciably better for those patients with negotiated endings. Hence, a key target for these – and any – service is to keep patients in treatment sufficiently long to achieve an agreed ending and completed course of treatment.

5.1.28   Benchmarks

Second, in relation to our use of benchmarks, we used a combination of data sources and published literature in order to place the current IAPT outcomes in historical context. In terms of the former (i.e., data sources), the availability of the CORE data set provided an historical context and while the patient sample selected according to severity and duration appeared comparable to those in the demonstration sites, it should be borne in mind that the data were collected by services without DH funding and not resourced at that time to collect sessional measures. Further, it is likely that condition-specific measures (i.e., PHQ-9 and GAD-7) and generic outcome measures (i.e. CORE-OM) may differ in their sensitivity to change. However, we found the CORE-OM to be conservative (i.e. more stringent) in defining recovery when compared with the combined PHQ-9/GAD-7 criterion regardless of the GAD-7 cut-off adopted. The body of studies included in the trials literature is likely to grow apace, particularly in relation to the self-help and computer-supported interventions. In this respect, it is likely that more focused benchmarks will be available in future that may enable a more precise location of the IAPT outcomes in relation to the trials literature.

6 Cohort study of comparative costs, psychological health, general health and well-being

Overview

In this chapter, the costs and outcomes of patients recruited from General Practitioners (GP) in the two Improving Access to Psychological Therapies (IAPT) demonstration sites are compared to the costs and outcomes of patients recruited from GP practices in matched comparator sites. The aim was to assess the cost effectiveness of the new delivery mode of providing improved access to Cognitive Behavioural Therapy (CBT). A cohort of patients was recruited via GP sites in Doncaster and Newham as well as matched sites in Wakefield, Barnsley, and City and Hackney. Outcomes and service use data were collected from these patients (self-report) at baseline and at two follow-up points, 4 months and 8 months. IAPT cost and service provision information was provided by the two demonstration sites. Cost effectiveness was assessed based on this information.

Methods

6.1.1 Study design

The overall study was designed to test one central hypothesis: that access to primary care CBT psychological therapies via newly configured service delivery models implemented at two demonstration sites would be cost effective when compared with services at existing comparator sites. The health economic component of the study analysed the consequences of the new CBT investments at the two demonstration sites in terms of costs, psychological and general health related quality of life and overall well-being. A comparative cohort design was adopted in order to overcome the potential biases and problems with using uncontrolled routine data on patients attending IAPT.

The advantages of this study design were:

- There was a risk of bias from the selection of patients going to IAPT making them different from the cases in primary care in general that may have made them more amenable to improvement. Our design aimed to avoid this problem by recruiting cohorts of patients from primary care in the demonstration site prior to referral to IAPT.

- IAPT was not just an individual level intervention and should have been seen as impacting on the population as a whole. The benefits of IAPT may have been experienced by those not using the IAPT, such as from resources released in the practice to use on others. This was captured by the patients in the cohort in the demonstration sites who did not use IAPT.

- An uncontrolled design would not have provided a sound basis for estimating the outcome from IAPT since the patients may have improved under usual care without IAPT. This study therefore aimed to recruit patients from matched control sites in order to estimate the incremental improvement in outcomes.

6.1.2 Recruitment

6.2.2.1 Sites

The demonstration sites in Doncaster and Newham were compared to other sites of similar socio-demography and comparable local changes. Data held on the National PCT Database housed at the National Primary Care Research & Development Centre (NPCRDC) at the University of Manchester, was interrogated and a strategy of statistical 'nearest neighbours' was used to identify five PCTs nationally which most closely matched each of the two demonstration sites. The variables used in the analysis were: multiple index of deprivation, % white ethnicity, % population under 30, average list size, and average QOF points. Within the 'nearest neighbour' set, comparators were also matched on a) geographical location b) local implementation of 'pathways to work', c) recent changes in PCT organisational structure and d) ethnic diversity. Sites within this pool of PCTs were approached to gain additional information in relation to service configuration, capacity, and case mix so that the most appropriate comparator sites could be selected.

6.2.2.2 General practices

The original intention was to sample clusters of GP practices using a random sample stratified by the key characteristics listed above. A pool of practices was to be identified and approached, with the aim being to recruit six practices within each demonstration and comparator site. However, this plan had to be modified as the difficulties of recruitment started to become clear.

In Doncaster, Wakefield and Barnsley, every GP practice within each site was approached in writing (Appendix 6.1) and invited to take part in the cohort recruitment. The Newham service had a phased approach to GP referrals, only those practices engaged with and referring to the IAPT

service (n=24) were approached. A second letter was sent out (Appendix 6.1) to encourage GPs to take part. GPs that consented completed a consent form (Appendix 6.2).

The initial stage of recruitment of GP practices from the London comparator sites was carried out by the Primary Care Research Network - Greater London (PCRN-GL). Practices were sent a letter with a response slip to indicate level of interest and this was followed up by a telephone call to the practice manager. An article was also placed in the PCRN-GL newsletter which was received by all Tower Hamlets and City and Hackney practices. Seven practices from City and Hackney and five from Tower Hamlets expressed an interest or asked for further information. These were followed up by a researcher from Sheffield University. Five practices were successfully recruited from City and Hackney. From Tower Hamlets two single handed practices agreed to take part in the study but no patients were recruited in the first three months. Of the remaining three practices, only one expressed an interest in taking part. It was therefore decided it would be unethical to recruit from this one practice as the numbers required for the study would not be reached. Hence, whilst nearly half the practices approached in Doncaster and Newham agreed to take part, a considerably smaller proportion in the comparator sites consented. Table 25 summarises the position.

Table 25. GP recruitment

Site	Approached practices	Agreed to take part, and sent out recruitment packs to patients	Successfully recruited patients
Doncaster	46	22	18
Barnsley	43	15	8
Wakefield	42	13	3
Newham	24	12	9
City & Hackney	42	5	5
Tower Hamlets	43	2	0
Total	**240**	**69**	**43**

6.2.2.3 Patients

Each practice which consented to take part was asked to prospectively identify patients who met the study criteria, or who had been seen by GPs within the last 4 weeks. Individual GPs within the practice were given a recruitment record sheet (Appendix 6.3), with the following criteria:

'*Anyone of working age (16 -64 years), newly presenting (or re-presenting) with anxiety or depression, who may benefit from a 'talking treatment' (any community based counselling or psychological therapy service)[30][31]*'

The same criteria were used in all sites. GP practices were also provided with posters to encourage recruitment (Appendix 6.4).

The practices mailed out an initial recruitment pack to patients shortly after they had been seen and recorded by the GP (or identified from the practice information system), and then a reminder letter approximately two weeks later. The recruitment pack contained a covering letter from the GP practice, an invitation letter and information sheet from the University of Sheffield, a copy of the baseline questionnaire, a consent form, a prepaid response envelope and the offer of a £10 gift voucher for returning the questionnaire (see Appendix 6.5-6.8). All responses were mailed directly to the University of Sheffield by the patient, and the patient's consent, contact details and responses were entered in to separate secure databases. The contact information was then used to re-contact patients at the four and eight month follow up stage. Everyone who responded at each stage was sent a £10 high street gift voucher. For those who responded to all 3 stages, a total of £30 worth of vouchers was sent. A reminder letter (Appendix 6.9) was sent out to potential participants by GPs two weeks after the initial invitation.

Recruiting was supposed to occur before patients were seen in IAPT. However, patients were responsible for sending back questionnaires and there were therefore delays. As a result, 76 IAPT patients (50 in Doncaster and 26 in Newham) were recruited into the study who had already had at least one session with the IAPT service when they sent back their questionnaires. In some cases, patients were re-referrals due to the long period of recruiting into the study. This problem was only identified when the electronic download data was made available at the end of the study period. These patients were retained in the analysis due to small sample sizes, but may have consequences for baseline assessments of outcomes and service use.

---

[30] I.e., excluding only: high complexity, high impairment, high substance use, high risk or high number of previous interventions with unhelpful outcomes.

[31] even if talking treatments are not available locally or the patient is offered and declines

6.1.3 Patient reported outcome measures (PROMs)

We adopted the patient reported outcome measures (PROMS) from the IAPT evaluation framework of the PHQ-9, GAD-7 and CORE-OM (though this was later dropped from the IAPT routine data). Additional outcome measures used in this study were the SF-6D, items to assess general well-being taken from the British Panel Household Survey and items concerned with employment. These measures were administered by postal questionnaire (Appendix 6.8) to all individuals in the demonstration and comparative cohorts.

For the cost effectiveness comparison, the primary measure was the Short Form 6 dimensions (SF-6D). The SF-6D is a preference-based generic measure of health designed for calculating quality adjusted life years (QALYs). It is composed of six multi-level dimensions of health: physical functioning, role limitation, social functioning, bodily pain, mental health and vitality. It was constructed from the Short Form 36 (SF-36) and valued using the standard gamble (SG) valuation technique on a sample of 611 members of the UK general population [71]. The resultant algorithm can be used to convert SF-36 data at the individual level into a preference-based index. The SF-6D has been chosen because the original SF-36 has been found to perform well in populations with common mental disorders and at the same time it meets the NICE reference case for use in economic evaluations of health care interventions at the time the study was designed [72].

General well-being measures consisted of two satisfaction and two happiness questions taken from the British Household Panel Survey (BHPS). The satisfaction questions covered overall life satisfaction and satisfaction with achievements with scores ranging from 0 (completely dissatisfied) to 7 (completely satisfied). The happiness questions covered frequency of happiness from 1 (none of the time) to 5 (all of the time) and feelings from 1 (so unhappy life is not worthwhile) to 5 (happy and interested in life).

The self-completed questionnaire included items about current employment and any time off work. An additional question that was not part of the SF-6D was included that focused on limitations in work and regular activities caused by emotional health.

6.1.4 Resource use

Resource use at baseline was obtained from the self-completed questionnaire for all cohort patients (Appendix 6.8). The questionnaire was adapted for use from other studies that have been conducted in the School of Health and Related Research (ScHARR) at the University of Sheffield. It consisted of two sections, one related to mental health service use which

asked about the number of times mental health professionals (viz. psychiatrist, psychologist, community psychiatric nurse psychotherapist/counsellor, other mental health professionals and voluntary sector services) were seen by respondents in the previous 4 months. The second section related to other primary and secondary health care and personal social service use. This questionnaire was also administered at 4 and 8 months.

For IAPT patients, details of IAPT service use in terms of time and sessions were obtained from each site in the form of an electronic download. This was done at the end of the study period in April 2010. This download was supplemented with data from the larger download used in Chapter 5.

In the original design, medication data was to be collected directly from medical records in primary care. However, resource limitations meant that this was not possible and the only available information was from the GP download data described in Chapter 7 for all IAPT patients. The cohort patients' medication data could not be matched with the study period time points and was only available for the demonstration sites, so these data were not suitable for use in this economic evaluation. However, no differences in medication use were found between IAPT and control cases in the IAPT data set reported in Chapter 7 of this report.

6.1.5 Sample size

The sample size calculation for this study was based on the potential health economic importance of the findings from this study separately for each demonstration site. An incremental cost for the service of £1,000 per patient, and an acceptability threshold of £20,000 per quality adjusted life year (QALY) (NICE, 2004), was assumed which resulted in required difference of 0.05 QALYs over a year. For a 5% two-sided level of significance and 80% power, assuming a difference between demonstration site and control of 0.05 QALY and a standard deviation of 0.15 QALY then a sample size of 142 per group (284 between a demo site and two comparators) was required. The data were actually clustered by general practice, so an adjustment was required for this. The intra class correlation was assumed to be comparable with that found for other general practice populations of 0.002 to 0.00818 and an average cluster size of 30 patients per practice. Then the total sample size across two demo sites plus four comparators was inflated to between 600 and 700. Allowing for a rate of refusal & attrition of 40% required a total N of 1440 broken down as follows: Demo A=360, A1=180, A2=180, Demo B=360, B1=180, B2=180.

6.1.6 Analysis of health, well-being and resource use

The primary comparison of costs and outcomes was between each IAPT site against its own control (i.e. Doncaster versus Wakefield & Barnsley; Newham versus City & Hackney). To combine results for the IAPT sites could have distorted the results due to differences between the services provided at the two demonstration sites and the fact that the comparators were matched for one IAPT site and not the other. However, to assist with the interpretation of the results and to allow comparison with the benchmarking presented in Chapter 5, a secondary comparison was undertaken between those seen, those not seen in the demonstration sites and the comparator cases.

Parametric (Independent Samples T Test; One Way Analysis of Variance) and non-parametric (Fisher's Exact Test; Chi Square Test; Mann-Whitney U test; Kruskal Wallis test) tests were undertaken to compare socio-demographic, health, well-being and resource use at baseline between sites and those who were followed up to those who were not followed up at 4 and 8 months. Follow-up data were compared in terms of changes in health, well-being and resource use using simple parametric (Paired Samples T test) and non-parametric tests (Wilcoxon Signed Rank Test; McNemar Test). Ordinary least squares (OLS) regression analysis was also undertaken that identified changes in the health and mental health for the different sites when controlling for age, gender and baseline levels of health related quality of life and mental health.

NHS Psychotherapy/counsellor use for those individuals who had been recruited after they had been seen in the IAPT service was adjusted using data from the services. In particular, baseline reported levels were adjusted downwards while 4 month psychotherapy use was adjusted upwards.

It was important to compare whether cohort study scores were similar to the rest of the IAPT service scores reported in Chapter 5 as any biases in the cohort would have implications on the transferability of findings reported in this Chapter. This was done by comparing PHQ 9 baseline and follow-up scores from the cohort study in Doncaster as the sample size was larger to PHQ 9 first and last scores (intention to treat) reported from the IAPT services download data reported in Chapter 5 for Doncaster. Comparison of self-reported resource use and IAPT download data resource use was also carried out.

6.1.7 Costing resource use

The costing took a National Health Service (NHS) and Personal Social Services (PSS) perspective and so covered all available health and social services resources, including the new CBT service itself and the

consequences for use of primary, community and secondary health care, and social services.

The IAPT services were costed based on financial data and service provision at the demonstration sites for the duration of the study (2007-2009). Financial data were obtained from the Trusts managing the services for the full costs of running the services for the two years of the study including training, preparation, equipment, facilities, and overheads to provide precise estimates of the intervention cost. Service provision information was extracted from electronic records provided by the IAPT services which included details of the amount of time and the number of sessions provided to patients. The total cost data over the 2 years were combined with total activity data from the services over the same period and average cost per minute was calculated for the two demonstration sites. This cost per minute was then used to cost the contact time of the cohort IAPT cases.

Use of non IAPT health and social care services was costed using published national unit cost data for the year price 2008/9 [73] and NHS Reference costs 2008/09 [74] as well as from evaluations uprated to 2008/9 prices [75]. All health and social care services in the comparator sites were also costed using national unit costs.

The primary analysis took an NHS and PSS perspective but a broader perspective of costs was also taken by the inclusion of productivity impact by costing the lost number of days from work. A human capital approach was used where the cost of each day of lost employment was assumed to be equal to the age and gender-specific national median daily wage rates. Full or part time status was also taken into account with differing wage rates applied for the two groups. 2008/2009 wage rates were used [76].

6.1.8 Economic evaluation

The economic evaluation included an assessment of the cost effectiveness of the intervention from a NHS and PSS perspective in terms of their incremental cost per quality adjusted life year (QALY). The economic evaluation was undertaken primarily using data from the individual level cohort data from the IAPT demonstration sites and the matched control sites. The SF-6D preference-based measure at baseline and the two follow-ups provided a basis for estimating a difference in outcome that was used to estimate the QALY gain (or loss) experienced by estimating the area under the curve over 8 months minus the baseline for each patient in the cohorts. Cost effectiveness in the primary analysis was assessed in terms of the ratio of the mean incremental cost per QALY for Doncaster and Newham compared to their controls and in the secondary analysis for patients seen in IAPT compared to the comparator site.

Productivity costs were calculated separately but they were not included in the incremental cost per QALY analysis.

6.1.9 Sensitivity analysis

Probability sensitivity analysis (PSA) was undertaken to examine the uncertainty surrounding the cost effectiveness ratios. Estimates of QALY gains and cost per patient were bootstrapped with 1000 replications and presented as cost-effectiveness acceptability curves (CEACs) which showed the probability of IAPT being cost effective. This was done for both primary and secondary analysis as well as for the EQ-5D estimates discussed below.

Any comparison to the NICE ICER threshold was difficult since we used SF-6D to calculate QALYs, whereas NICE prefers to use EQ-5D and these two measures have been shown to generate different values. In order to improve comparability with the NICE metric, SF-6D was translated into EQ-5D using an empirical mapping function. A separate dataset with both self-reported SF-6D and EQ-5D was used to estimate this function[32]. These predicted EQ-5D scores were used to calculate QALY gains (or losses) and the incremental cost per QALY which were compared to the SF-6D ICERs. PSA was also carried out using the predicted EQ-5D scores.

Sensitivity analysis of the IAPT costs was carried out by using national unit costs in place of IAPT unit costs to evaluate the effect of different costs on the cost effectiveness of the service.

Results

6.1.10 Recruitment

6.3.1.1 Sites

Wakefield and Barnsley were selected as the matched comparator sites for Doncaster, and City and Hackney and Tower Hamlets were selected as the matched comparator site for Newham. However, as discussed in section 6.1.2 Tower Hamlets was dropped due to the poor response from GPs. In addition, City and Hackney became a Wave 1 IAPT site during the study which has implications on outcomes and service use of patients recruited in this site.

---

[32] The function used is: EQ-5D = 1.563572*SF-6Dscore - 0.3502361. Results were capped at 1.0 for the highest values.

6.3.1.2 GPs

A total of 197[33] GP surgeries were approached and of these 67 (34%) agreed to take part in the study (Table 25). 43 (64%) of the practices that were recruited were successful in recruiting patients into the study. The consort diagram (Figure 13) details the numbers eligible for this study through the packs sent out by the 67 GP practices.

6.3.1.3 Patients

There were 5142 packs sent to GPs and out of these 4399[34] were handed out by the GP to patients who met the eligibility criteria. In all there were 529 returned by patients, representing an overall response rate of 12.0%. Response rates by site were as follows: Doncaster - 14.4%; Wakefield and Barnsley - 8.2%; Newham -12.8%; City and Hackney- 12.1% .

Of the 529 recruited, distribution across the sites was as follows: Doncaster 54.6%, Wakefield and Barnsley 21.6%, Newham 14.4%, and City & Hackney 17.3% (Figure 13). One individual died and one completed the questionnaire with reference to a study period outside the 4 month period (both in Doncaster) which resulted in a sample size of 527 at baseline. Of the 527, 4.4% (n=23) answered a short version of the questionnaire which contained only the SF-6D as well as the well-being questions. These were in Newham (n=6) and City and Hackney (n=17). As noted, 76 patients had had IAPT intervention and they were retained in the analysis.

At 4 months, 435 out of 528[35] responded resulting in an attrition rate of 17.6% (n= 93). Of those who responded, 3.6% (n= 19) answered the short version of the questionnaire which contained only the SF-6D and well-being questions.

At 8 months, a total of 425 out of 479[36] study participants responded representing an attrition rate of 2.3% (n = 10) from the 4 month stage giving an overall study attrition rate of 19.4% (n = 102) from baseline to the 8 month time point. 19 respondents answered the short version of the questionnaire at 8 months.

---

[33] Excluding Tower Hamlets

[34] Based on an estimate from GPs that sent back information on the number of packs given to patients.

[35] 528 includes the 527 who returned baseline questionnaires plus one incorrect one

[36] 479 was made up of the 435 who responded at 4 months plus 44 who responded at baseline, but not at 4 months.

At all stages, the analysis was restricted to the sample of patients who complete the full version of the questionnaire: baseline - 504; 4 months - 416; 8 months - 406.

6.3.1.4 IAPT seen patients

154 and 62 referrals of patients that had also taken part in the cohort study were identified in the Doncaster and Newham IAPT service datasets at the end of the study (April 2010). The 62 referrals in Newham were from 53 unique patients whereas the referrals in Doncaster represented unique patients. This included both patients who had been referred and seen and those who had been referred but who had not been seen in the IAPT services.

12/154 in Doncaster were referred but had no record that they were seen and 10/62 referrals in Newham[37] from 5/53 individuals were referred but not seen. There were therefore 142 and 48 patients who were seen in the IAPT services in Doncaster and Newham respectively. Secondary analysis focuses on these patients who were seen compared to those who were not seen and those in the comparator sites.

---

[37] 7 were recorded as not seen/no contact/ failure to engage and 3 were identified as step1 assessments but there was no contact time or mental health scores recorded and these were treated as not seen.

Figure 13.   Consort diagram for GP cohort study
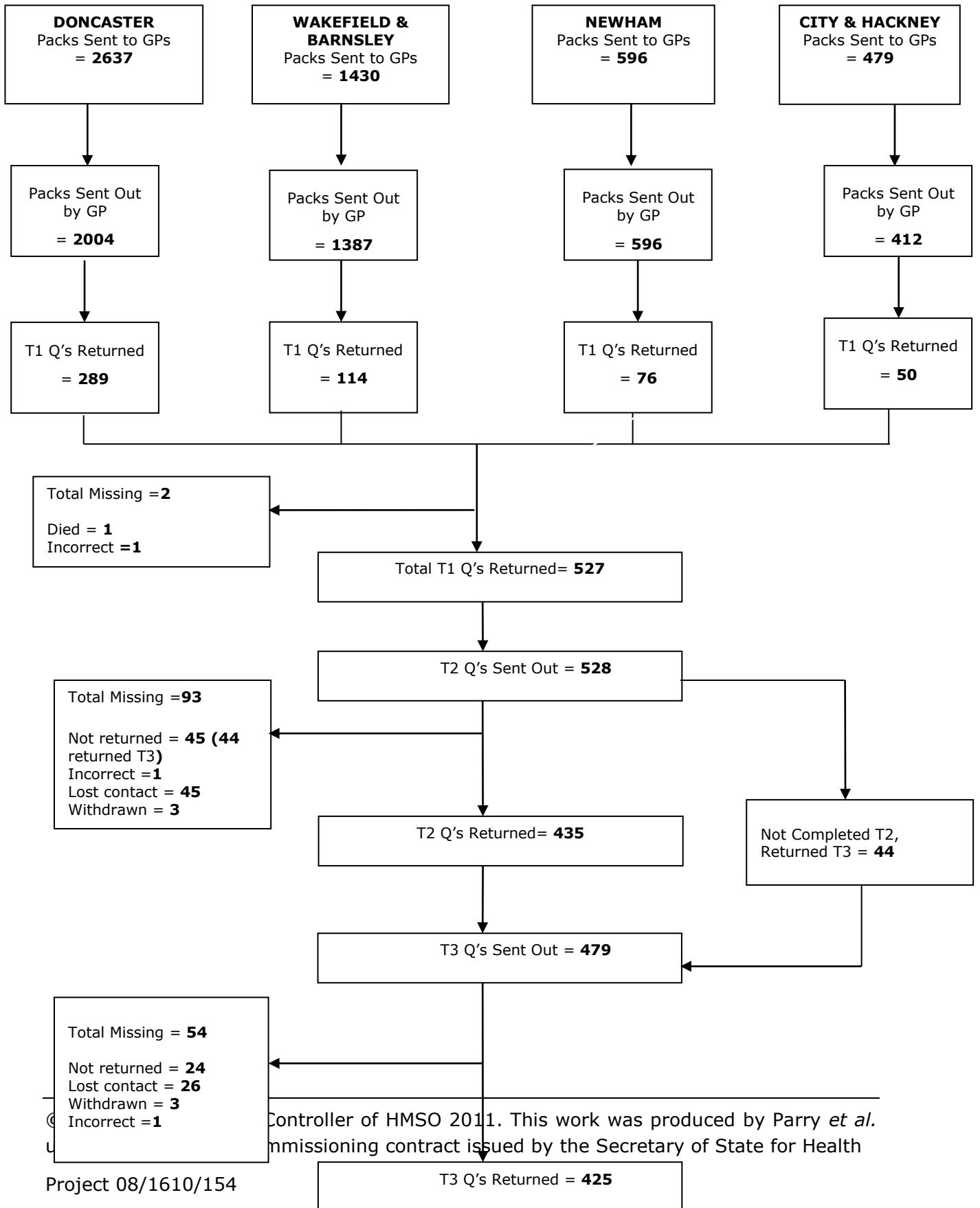
DONCASTER
Packs Sent to GPs
= **2637**

**WAKEFIELD &
BARNSLEY**
Packs Sent to GPs
= **1430**

NEWHAM
Packs Sent to GPs
= **596**

CITY & HACKNEY
Packs Sent to GPs
= **479**

Packs Sent Out
by GP
= **2004**

Packs Sent Out
by GP
= **1387**

Packs Sent Out
by GP
= **596**

Packs Sent Out
by GP
= **412**

T1 Q's Returned
= **289**

T1 Q's Returned
= **114**

T1 Q's Returned
= **76**

T1 Q's Returned
= **50**

Total Missing =**2**

Died = **1**
Incorrect **=1**

Total T1 Q's Returned= **527**

T2 Q's Sent Out = **528**

Total Missing =**93**

Not returned = **45 (44**
returned T3**)**
Incorrect =**1**
Lost contact = **45**
Withdrawn = **3**

T2 Q's Returned= **435**

Not Completed T2,
Returned T3 = **44**

T3 Q's Sent Out = **479**

Total Missing = **54**

Not returned = **24**
Lost contact = **26**
Withdrawn = **3**
Incorrect =**1**

Project 08/1610/154

T3 Q's Returned = **425**

6.1.11   Questionnaire completion rates

At baseline, PROMs completion rates range from 95.4% (481/504) in the SF-6D to 99.6% (502/504) in the happiness question.

The completion rate for the SF-6D at 4 months was 97.8% (407/416). Combining this with the baseline completion rates results in a sample size of 396 (95.2%) who have both baseline and 4 months SF-6D scores.

Completion rates for the SF-6D at 8 months were 96.1% (390/406). Combining this with the baseline completion rates results in a sample size of 378 (93.1%) who have both baseline and 8 months SF-6D scores.

6.1.12   Baseline

6.3.3.1 Demographics

Of the 504 respondents[38] who answered the full questionnaire at baseline, the majority were female (71.4%) (Table 26). Within the sites, Doncaster, and Wakefield and Barnsley had a majority of females (72% and 75% respectively) with no significant difference. However, in Newham, 59% were female compared to 81% in City and Hackney (Fisher's Exact Test, p<0.001). The mean age across the whole cohort was 40.9 (sd = 14.3) with no statistically significant differences across the IAPT and comparator matched sites.

In Doncaster, Wakefield and Barnsley, over 96% of respondents reported their ethnicity as white whereas in Newham and City and Hackney this was 53% and 67% respectively (Table 26). In Newham, 24% reported their ethnicity as Asian[39] and 9% Black/Black British. In City and Hackney, 6% were Asian[40] and Black/Black British were 6% with these distributions significantly different from those in Newham (Chi Square = $9.2_{(3)}$, p<0.05).

Patients who were identified as IAPT seen patients from the IAPT demonstration sites were not significantly different from not-seen patients or patients in the comparator sites in terms of gender, age and ethnicity.

---

[38] As noted in the previous section, 23 individuals answered a short version of the questionnaire. Those who answered the short version of the questionnaire were not significantly different in terms of their SF-6D and well-being scores but they reported better mental health (SF-6D dimension) than those who answered the full questionnaire (Mann Whitney U = 4457.5, Z = -1.97, p = 0.048). To allow comparability across the different measures including the mental health measures and resource use, those who answered the short version (n=23) were excluded from the rest of the analysis. Demographics for the full-sample were reported in Appendix 6 Table 1.

[39] Pakistani 12%, Indian, Bangladeshi and other Asian 12%

[40] Pakistani 3%, Bangladeshi 3%

When seen patients were compared to not-seen patients within the IAPT sites there were no differences by gender and ethnicity but there was some evidence that those who were not seen were more likely to be older (>39 years) (Fisher's Exact Test, $p<0.1$) compared to those who were seen in IAPT.

Table 26. Age, Gender and Ethnicity of Cohort Study Patients

	All		Doncaster		Wakefield & Barnsley		Newham		City and Hackney	
	n	%	n	%	n	%	n	%	n	%
Females	360	71.4	207	72.1	85	74.6	41	58.6	27	81.2
Age (sd)	40.9 (14.3)		41.7 (14.9)		40.5 (13.5)		38.5 (12.2)		42.8 (13.6)	
Ethnicity										
Asian/Asian-British	19	3.8	0	0.0	0	0.0	17	24.3	2	6.1
Black / Black-British	9	1.8	1	0.3	0	0.0	6	8.6	2	6.1
Chinese	0	0	0	0.0	0	0.0	0	0.0	0	0.0
White	450	89.3	278	96.9	113	99.4	37	52.9	22	66.7
Mixed	5	1.0	2	0.7	0	0.0	2	2.8	1	3.0
Other	9	1.8	0	0.0	0	0.0	5	7.1	4	12.1
No ethnicity reported	12	2.4	6	2.1	1	0.6	3	4.3	2	6.1
Total	504	100	287	100	114	100	70	100	33	100

6.3.3.2 Patient Reported Outcome Measures (PROMs)

Primary analysis

The mean SF-6D score across the whole cohort was 0.61 (sd = 0.13) with a minimum of 0.30 and a maximum of 0.96 (Table 27). The mean (sd) score for Wakefield and Barnsley was 0.63 (0.13), which was higher than Doncaster at 0.61 (0.13) but this was not a statistically significant difference. Both Newham and its matched site had a mean SF-6D score of 0.61. These SF-6D scores are substantially lower than reported general population scores of 0.80 (sd = 0.12) [77].

The mean life satisfaction scores were 2.77 (sd = 1.77; median = 3; mode =3) for overall life and 2.59 (sd = 1.88; median = 2; mode = 1) for satisfaction with achievements. Happiness levels had a mean of 2.42 (sd = 0.95; median =2; mode = 2) for frequency and 2.73 (sd = 1.06; median = 3; mode =3) for feelings. Wakefield and Barnsley had statistically significant higher baseline happiness levels than Doncaster in terms of both frequency (p <0.1) and feelings (p <0.05). There were no significant differences across the other well-being measures (Table 27).

The mean (sd) PHQ 9 scores were 15.7 (7.1) across the cohort. There was some evidence that Wakefield and Barnsley had lower PHQ 9 scores 15.1 (7.2) compared to Doncaster 16.4 (7.1) (p <0.1) that indicated better psychological health. Mean (sd) scores for Newham 13.7 (7.1) were lower than those for City and Hackney 14.8 (5.6) but this difference was not statistically significant. PHQ 9 severity levels for the cohort indicated that 77% of the respondents fell in the clinical population, ranging from 69% to 81% within the IAPT and comparator sites (Table 28 and Appendix 6 Table 6.2).

Mean (sd) GAD 7 scores were 13.4 (5.7) across the cohort with no significant differences across the matched sites (Table 27). Sixty nine percent of the respondents fell in the clinical population, ranging from 60% to 76% within the IAPT and comparator sites (Table 28 and Appendix 6 Table 6.3).

Mean (sd) CORE OM were 20.1 (7.8), and again there were no significant differences across the matched sites (Table 27)[41]. 87% of the respondents

---

[41] 12 individuals had 4 or more missing items in the CORE and they had significantly lower SF-6D scores (0.53 vs. 0.61, ($t_{(7.3)}$= 2.2, p<0.1) and they were also older (58.7 vs. 40.5, ($t_{(9.3)}$= -3.8, p<0.01).

had CORE-OM scores that indicated they were in the clinical population (Table 28 and Appendix 6 Table 6.4).

All the health and well-being scores indicated that cohort patients in Doncaster were slightly worse off than the other sites although this was not a statistically significant difference.

Secondary analysis

Those who were seen in IAPT services had a slightly lower mean (sd) SF-6D scores of 0.60 (0.13) than the not-seen patients of 0.62 (0.14) or those in comparator sites of 0.62 (0.12). However, these were not statistically significant differences (Appendix 6 Table 6.5).

Well-being and happiness measures were similar across those who were seen in IAPT services and those who were not seen as well as those in the comparator sites (Appendix 6 Table 6.5).

Mean psychological health scores were slightly lower for the comparator sites than both the IAPT seen patients and not-seen patients (Appendix 6 Table 6.5). Mean (sd) PHQ 9 scores in the comparator site were 15.5 (6.9) compared to IAPT seen patients 15.8 (7.1) and not-seen patients 16.1 (7.3). Corresponding GAD 7 mean (sd) scores for the comparator sites were 13.1 (5.6) compared to IAPT seen patients 13.6 (5.5) and not-seen patients 13.3 (5.9) whereas CORE-OM scores were 19.7 (7.8) compared to IAPT seen patients 20.2 (7.8) and not-seen patients 20.4 (8.0) (Appendix 6 Table 6.5). However, all these differences were small and none of them were statistically significant.

Table 27. Baseline Patient Reported Outcomes Measures: Health and Well-being

		All		Doncaster		Wakefield and Barnsley	Test[a]		Newham		City and Hackney	Test[a]
	n	Mean (sd)	n	Mean (sd)	n	Mean (sd)	P Value	n	Mean (sd)	n	Mean (sd)	P Value
SF-6D	481	0.61 (0.13)	271	0.61 (0.13)	112	0.63 (0.13)	0.12[b]	67	0.61 (0.15)	31	0.61 (0.13)	0.97[b]
Life satisfaction	500	2.77 (1.77)	284	2.63 (1.71)	114	2.84 (1.86)	0.24	69	3.12 (1.80)	33	3.03 (1.83)	0.75
Satisfaction achievements	496	2.59 (1.88)	281	2.44 (1.88)	113	2.72 (1.88)	0.16	70	2.91 (1.83)	32	2.72 (1.89)	0.57
Happy (feelings)	500	2.73 (1.06)	284	2.64 (1.02)	114	2.84 (1.05)	0.09	69	2.87 (1.21)	33	2.73 (1.10)	0.55
Happy (frequency)	502	2.42 (0.95)	285	2.33 (0.93)	114	2.56 (1.02)	0.04	70	2.51 (0.93)	33	2.55 (0.90)	0.86
PHQ9	500	15.7 (7.1)	285	16.4 (7.1)	113	15.1 (7.2)	0.10	69	13.7 (7.1)	33	14.8 (5.6)	0.45
GAD7	500	13.4 (5.7)	284	13.8 (5.7)	113	13.3 (5.8)	0.39	70	11.9 (5.5)	33	12.6 (4.6)	0.48
CORE	492	20.1 (7.8)	280	20.5 (7.9)	111	19.6 (8.1)	0.33[b]	68	19.0 (7.7)	33	20.0 (6.8)	0.52[b]

504	287	114	70	33

a Mann Whitney test comparing matched sites (Doncaster vs. Wakefield and Barnsley; Newham vs. City and Hackney) PROMS

b Independent Samples T Test comparing matched sites PROMS

Table 28. Baseline severity levels in PHQ 9, GAD 7 and CORE-OM

	All		Doncaster		Barnsley and Wakefield		Newham		City and Hackney	
	n	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**Valid %**
PHQ-9 cut-off=10										
Non-clinical population	110	21.8	53	18.5	29	25.4	21	30.0	7	21.2
Clinical Population	390	77.4	232	80.8	84	73.7	48	68.6	26	78.8
Missing	4	0.8	2	0.7	1	0.9	1	1.4	0	0
GAD 7 cut-off=10										
Non-clinical population	150	29.8	78	27.2	36	31.6	28	40.0	8	24.2
Clinical Population	350	69.4	206	71.8	77	67.5	42	60.0	25	75.8
Missing	4	0.8	3	1.0	1	0.9	0	0	0	0
CORE-OM cut-off=10										

Project 08/1610/154

Non-clinical population	53	10.5	29	10.1	13	11.4	8	11.4	3	9.1
Clinical Population	439	87.1	251	87.5	98	86.0	60	85.7	30	90.9
Missing	12	2.4	7	2.4	3	2.6	2	2.9	0	0
Total	**504**		**287**		**114**		**70**		**33**	

6.3.3.3 Employment

Primary analysis

256/504 (51%) of the patients in the cohort were in full time or part-time employment, ranging from 44% to 53% within the sites (Table 29). 144/504 (29%) were unemployed. This figure was substantially higher than national unemployment rates which ranged between 5.2% and 7.9% in England in 2008-2009 [78]. There were no statistically significant differences in employment activity when comparing the IAPT sites and their matched comparator sites.

161/256 (63%) of those who were in employment reported days absent from work with a mean (sd) of 13.2 (24.0) days (Table 29). The median was 2.5 days for the whole sample and Doncaster has the highest median at 4 days. There was a significant difference between days absent reported in Doncaster compared to Wakefield and Barnsley (Mann Witney U = 3552 Z = -1.99, p = 0.047), but there were no significant differences between Newham and City & Hackney.

Of the 205 who answered the sick pay question, 134 (65%) were receiving sick pay and there were no statistically differences in the matched sites.

146/504 (29%) report that they were receiving benefits with no statistical differences across the matched sites.

44.3% (222/501) of the patients in the cohort study reported that their emotional health affected their work or regular activities most/all of the time with similar proportions across the sites apart from in Newham where this 37.1%.

Secondary Analysis

Patients who were identified as IAPT seen patients from the IAPT demonstration sites were not significantly different from not-seen patients or patients in the comparator sites in terms employment, lost days or benefits status (Appendix 6 Table 6.6).

Table 29. Employment Status

		All	Doncaster	Wakefield and Barnsley	Newham	City and Hackney
		n (Valid %)	n (Valid %)	n (Valid %)	n (Valid %)	n (Valid %)
Employment Status	**Emp. F/T**	180(36.0)	99 (34.9)	41 (31.1)	25 (35.7)	15 (45.5)
	Emp. P/T	76 (15.2)	52 (18.3)	16 (13.9)	6 (8.6)	2 (6.1)
	Unemployed	144 (28.8)	76 (26.8)	35 (33.8)	23 (32.9)	10 (30.3)
	Student F/T	9 (1.8)	4 (1.4)	1 (1.1)	3 (4.3)	1 (3.0)
	Retired	44 (7.4)	23 (8.1)	8 (6.9)	4 (5.7)	2 (6.1)
	Homemaker / Carer F/T	37 (7.4)	24 (8.5)	12 (13.6)	6 (8.6)	2 (6.1)
	Other	10 (2.0)	6 (2.1)	0 (0.0)	3 (4.3)	1 (3.0)
Lost employment		161 (62.9)	101 (66.9)	32 (56.1)	17 (54.8)	11 (64.7)
Lost Days Mean (sd)		13.2 (24.0)	16.0 (25.0)	9.2 (21.8)	10.4 (26.0)	6.7 (14.3)
Lost Days Median		2.5	4.0	1.0	2.0	2.0
Receiving Sick Pay	**Yes**	134 (65.4)	85 (70.8)	27 (60.0)	14 (53.8)	8 (57.1)
	No	71 (34.6)	35 (29.2)	18 (40.0)	12 (46.2)	6 (42.9)
Receiving Benefits	**Yes**	146 (29.2)	77 (27.1)	35 (31.0)	24 (34.3)	10 (30.3)
	No	354 (70.8)	207 (72.9)	78 (69.0)	46 (65.7)	23 (69.7)
Emotional health limits regular work activities						
None of the time		32 (6.4)	19 (6.7)	8 (7.0)	4 (5.7)	1 (3.1)
A little of the time		99 (19.8)	52 (18.2)	23 (20.2)	18 (25.7)	6 (18.8)
Some of the time		148 (30.0)	77 (27.0)	37 (32.5)	22 (31.4)	12 (37.5)
Most of the time		144 (28.7)	93 (32.6)	24 (21.2)	15 (21.4)	12 (37.5)
All of the time		78 (15.6)	44 (15.4)	22 (19.3)	11 (15.7)	1 (3.1)

6.3.3.4 Resource use

Primary analysis

NHS Psychotherapist or counsellor use was the same for Doncaster versus Wakefield & Barnsley[42] at baseline (Appendix 6 Table 6.7). The use of other mental health services such as psychiatrists, community psychiatric nurses was small in both Doncaster and its matched sites with no significant differences.

Most individuals reported the use of primary care services, and in particular general practitioner (GP) services with 83% (239/287) having at least one contact in Doncaster and 74% (84/114) in Wakefield and Barnsley. The mean (sd) number of times individuals visited the GP surgery in Doncaster was 3.5 (6.1) and Wakefield and Barnsley 2.6 (3.3) which was a statistically significant difference ($p < 0.05$)[43].

The rest of the primary and secondary health care service use was not significantly different across Doncaster and its matched sites.

As in Doncaster and its matched sites, there were no statistically significant differences NHS psychotherapist or counsellor use at baseline in Newham compared to City and Hackney[44] (Appendix 6 Table 6.8). The use of other mental health services such as psychiatrists, community psychiatric nurses was small in both Newham and City & Hackney with no significant differences.

The proportion of individuals reporting GP use was 54% (38/70) and 85% (28/33) in Newham and City & Hackney respectively. Associated means (sd) were 2.2 (3.1) and 4.7 (12.1) for the two sites with a statistically significant difference between the two ($p = 0.06$)[45]. The rest of the primary and

---

[42] As noted, some of the study participants were recruited into the study while they were already receiving IAPT interventions. Download data was used to adjust the level of NHS psychotherapy use that was reported by patients in order to account for this.

[43] 2 individuals in Doncaster report very high levels of GP use, one individual reports seeing their GP 50 times while the other reports 80 contacts. Excluding these two outliers lowers the mean (sd) use in Doncaster to 2.9 (2.6) from 3.5 (6.1). This still represents a significantly higher use of GPs in Doncaster compared to the matched sites [Mann Whitney u = 13197 Z = -2.75 $p < 0.05$].

[44] As noted, some of the study participants were recruited into the study while they were already receiving IAPT interventions. Download data was used to adjust the level of NHS psychotherapy use that was reported by patients in order to account for this.

[45] 1 individuals in City & Hackney reports 70 contacts with a GP. Excluding this outlier lowers the mean (sd) use in City & Hackney to 2.6 (2.9). This still represents a significantly higher

secondary health care service use was not significantly different across Newham and City & Hackney.

Secondary analysis

11% (21/186)[46] of patients who were seen in the IAPT demonstration sites reported NHS Psychotherapist or counsellor use whereas this was 13% (23/171) for patients located in the IAPT sites but who were not seen by the service and 11% (16/147) in the comparator sites with no significant differences (Appendix 6 Table 6.9).

11% (20/186) of patients who were seen in the IAPT demonstration sites reported other unspecified primary health care service use whereas this was 6% (11/171) and 2% (3/147) for IAPT not seen and the comparator site patients. Associated mean (sd) were 0.6 (2.6) vs. 0.2 (0.8) and 0.03 (0.3) for the IAPT seen, IAPT not seen and comparator site groups respectively with significant differences between the groups (p = 0.01). The rest of the primary and secondary health care service use was not significantly different across the three groups.

6.1.13   Follow-up at 4 months

6.3.4.1 Followed up vs. not followed up

416/504[47] patients responded at the 4 month follow-up. Individuals who dropped out were younger (38.5 (13.1) vs. 41.3 (14.5), $t_{(123.9)}$ = -1.77, p = 0.079). 37.5% (33/88) of those who dropped out were men compared to 26.7% (111/416) for those who were followed up which was a statistically significant difference (Fisher's Exact Test p = 0.051). Of those not followed up 82.9% (68/82) were white compared to 93.2% (382/410) of those who were followed up (Chi Square $_{(3)}$ = 10.99, p = 0.012).

Patients who were not followed up did not differ significantly in terms of baseline SF-6D scores or most of the well-being scores compared to those who were. Those who dropped out were less happy (frequency) at baseline than those who were followed up (2.17 (0.92) vs. 2.48 (0.95); Mann Whitney U = 14950, z = -2.64, p = 0.008). They also reported higher core mean scores at baseline, 22.0 (sd = 7.60; n =85) compared to 19.7 (sd

---

use of GPs in City & Hackney compared to Newham [Mann Whitney u = 893 Z = -1.68 p =0.09].

[46] As noted, some of the study participants were recruited into the study while they were already receiving IAPT interventions. Download data was used to adjust the level of NHS psychotherapy use that was reported by patients in order to account for this.

[47] Those who answered the full version of the questionnaire.

=7.81; n = 409) and this difference was statistically significant ($t_{(490)}$ = 2.52, p = 0.012) but they did not have significantly different PHQ 9 or GAD 7 scores.

Patients who dropped out were not significantly different from those who were followed up in terms of employment status or benefits status. However, 52.8% (19/36) of those who dropped out were receiving sick pay compared to 68.0% (115/169) of those who were followed up (Fisher's Exact Test p = 0.062).

6.3.4.2 Patient Reported Outcomes Measures (PROMs)

Primary analysis

At 4 months, all PROMS scores indicated improvement in both the IAPT sites and their matched comparator sites (Table 30).

SF-6D preference-based scores increased by 0.041 in Doncaster and 0.026 in Wakefield and Barnsley which were statistically significant improvements. SF-6D scores also increased in Newham by 0.027 and in City and Hackney by 0.035 but these were not statistically significant increases. When these improvements were compared between the matched sites (Doncaster vs. Wakefield and Barnsley; Newham vs. City and Hackney), there were no statistically significant differences.

There were statistically significant improvements in all three measures of well-being at 4 months in Doncaster and its matched sites (P<0.05) (Table 30) and weak evidence of an improvement in happiness (frequency) (p<0.1). Comparison of these improvements by matched sites indicated that Doncaster had larger improvements compared to Wakefield and Barnsley in happiness (frequency). None of the differences between Newham and its control were statistically significant.

The PHQ 9, GAD 7 and CORE-OM showed statistically significant improvements overall, in Doncaster and its matched site as well as in City and Hackney. When changes were compared across the matched sites, the improvements in City and Hackney for the PHQ 9 and the GAD 7 were greater than those reported in Newham (p<0.05).

The multiple regression, after controlling age, gender and baseline scores site indicated no significant differences in change across any of the health or well-being variables except for 'happy feeling' where Doncaster was found to have a significantly larger gain (0.445, p<0.05) (Table 31). Matched IAPT and comparator sites results confirm these findings (Appendix 6 Tables 6.10 and 6.11).

Secondary analysis

PROM changes at 4 months indicated that there were statistically significant changes for IAPT seen patients, IAPT not seen patients and comparator sites in all the measures (Table 32). There were also statistically significant differences in the SF-6D changes when the 3 groups were compared, with the IAPT seen patients reporting higher improvements in SF-6D compared to both the IAPT not-seen patients and comparator sites (0.06 vs. 0.02 and 0.02, P=0.01) and frequency of happiness (0.52 vs. 0.25 and 0.25, P<0.001). After controlling for age, gender and baseline scores only the happiness differences remained significant (Table 33).

Table 30. Change in Health and Well-being 4m: IAPT versus comparator matched sites

	Doncaster		Wakefield & Barnsley		MW Test	Newham		City & Hackney		MW Test
	mean	sd	mean	sd	P value	mean	sd	mean	sd	P value
SF-6D	0.604	0.14	0.635	0.12	0.22[a]	0.594	0.15	0.617	0.12	0.81[a]
SF-6D 4m	0.645	0.15	0.661	0.14		0.621	0.15	0.651	0.11	
Change	0.041***	0.11	0.026**	0.10		0.027	0.13	0.035	0.10	
n	232		95			44		25		
Satisfaction	2.68	1.72	2.90	1.86	0.35	3.16	1.97	3.19	1.67	0.50
Satisfaction 4m	3.32	1.77	3.24	1.83		3.38	1.92	3.54	1.39	
Change	0.64***	1.52	0.34**	1.44		0.22	1.66	0.35	1.35	
n	242		100			45		26		
Happy (frequency)	2.38	0.93	2.66	1.01	0.04	2.51	0.97	2.62	0.90	0.95
Happy (frequency) 4m	2.80	0.97	2.82	0.91		2.78	1.04	2.73	0.92	
Change	0.42***	0.95	0.16**	0.91		0.27*	0.89	0.12	1.11	
n	242		100			45		26		
Happy (feelings)	2.65	1.01	2.89	1.02	0.67	2.73	1.34	2.88	1.03	0.73
Happy (feelings) 4m	3.25	1.10	3.43	1.07		3.00	1.31	3.04	0.96	
Change	0.60***	1.02	0.54***	0.98		0.27	1.37	0.15	0.97	
n	244		100			45		26		
PHQ-9	16.16	7.14	14.86	7.15	0.60	13.84	7.72	13.77	5.40	0.04
PHQ-9 4m	13.04	7.68	12.11	7.18		13.26	8.14	10.51	6.34	
Change	-3.12***	5.83	-2.75***	5.60		-0.58	5.21	-3.26**	4.80	
n	244		96			45		26		
GAD-7	13.63	5.77	13.09	5.82	0.94	12.21	5.98	12.33	4.90	0.04
GAD-7 4m	11.01	6.48	10.53	6.12		11.82	6.68	9.65	5.42	
Change	-2.61***	5.18	-2.56***	5.89		-0.39	4.32	-2.67**	4.39	
n	241		97			45		26		

	Doncaster		Wakefield & Barnsley		MW Test	Newham		City & Hackney		MW Test
	mean	sd	mean	sd	P value	mean	sd	mean	sd	P value
CORE -OM	20.11	7.91	19.13	7.83	0.14	18.95	8.65	18.63	6.35	0.29
CORE-OM 4m	16.40	8.77	16.41	8.29		18.04	8.93	15.92	8.09	
Change	-3.71***	5.31	-2.73***	5.76		-0.91	7.70	-2.71**	4.98	
n	237		95			44		26		

*** p<0.01, ** p<0.05, * p<0.1 Wilcoxon Signed Rank Test

MW Test – Mann-Whitney Test; a Independent Sample T-test

Table 31. Multivariate linear regression : Change in health and well-being from baseline to 4 months follow-up.

	SF-6D	PHQ 9	GAD 7	CORE OM	Life satisfaction	Happy (frequency)	Happy (feelings)
Doncaster	0.001	0.722	0.461	-0.824	0.117	0.192	0.346*
	(0.022)	(1.115)	(1.024)	(1.158)	(0.284)	(0.172)	(0.194)
Wakefield & Barnsley	-0.009	0.797	0.295	0.047	-0.077	0.054	0.378*
	(0.023)	(1.193)	(1.097)	(1.242)	(0.302)	(0.183)	(0.207)
Newham	-0.020	2.570*	2.078*	1.761	-0.083	0.120	0.069
	(0.026)	(1.338)	(1.230)	(1.397)	(0.340)	(0.206)	(0.233)
Baseline Score	-0.265***	-0.253***	-0.303***	-0.160***	-0.356***	-0.490***	-0.441***
	(0.040)	(0.038)	(0.043)	(0.036)	(0.038)	(0.043)	(0.044)
Age	-0.001***	0.008	-0.015	0.002	0.006	-0.003	-0.006*
	(0.000)	(0.019)	(0.017)	(0.020)	(0.005)	(0.003)	(0.003)
Sex	0.011	0.461	-0.309	0.598	0.046	-0.088	-0.116
	(0.012)	(0.606)	(0.561)	(0.639)	(0.155)	(0.093)	(0.106)
Constant	0.242***	-0.228	1.771	0.054	1.200***	1.556***	1.714***
	(0.038)	(1.449)	(1.355)	(1.583)	(0.356)	(0.233)	(0.261)
Observations	393	409	406	399	410	412	410
R-squared	0.11	0.12	0.12	0.07	0.19	0.26	0.22

Standard errors in parentheses

 *** p<0.01, ** p<0.05, * p<0.1

Reference category: *City & Hackney*

SF-6D, life satisfaction, happiness: positive coefficients indicate improvements over and above City and Hackney

PHQ-9, GAD 7, CORE-OM: negative coefficients indicate improvements over and above City and Hackney

Table 32. Change in Health and Well-being 4m: IAPT seen, IAPT not-seen patients and comparator sites

	IAPT seen		IAPT site not seen		Comparator site		Kruskal Wallis Test
	mean	sd	mean	sd	mean	sd	P value
SF-6D	0.59	0.13	0.62	0.15	0.63	0.12	0.01[a]
SF-6D 4m	0.64	0.15	0.64	0.15	0.66	0.13	
Change	0.06***	0.11	0.02**	0.11	0.03**	0.10	
n	148		128		120		
Satisfaction	2.77	1.76	2.74	1.79	2.96	1.82	0.36
Satisfaction 4m	3.34	1.77	3.31	1.82	3.30	1.75	
Change	0.57***	1.52	0.57***	1.58	0.34**	1.42	
n	155		132		126		
Happy (frequency)	2.34	0.89	2.47	0.99	2.65	0.98	0.00
Happy (frequency) 4m	2.86	0.98	2.72	0.98	2.80	0.91	
Change	0.52***	0.93	0.25***	0.95	0.15*	0.95	
n	155		134		126		
Happy (feelings)	2.66	1.03	2.68	1.12	2.89	1.02	0.31
Happy (feelings) 4m	3.28	1.17	3.14	1.11	3.35	1.05	
Change	0.62***	1.13	0.46***	1.03	0.46***	0.99	
n	153		133		126		
PHQ-9	15.93	7.12	15.65	7.45	14.63	6.81	0.48

	IAPT seen		IAPT site not seen		Comparator site		Kruskal Wallis Test
	mean	sd	mean	sd	mean	sd	P value
PHQ-9 4m	12.83	7.80	13.36	7.70	11.77	7.01	
Change	-3.10***	6.14	-2.29***	5.37	-2.86***	5.43	
n	155		134		123		
GAD-7	13.67	5.61	13.11	6.05	12.93	5.63	0.71
GAD-7 4m	11.23	6.72	11.04	6.28	10.34	5.97	
Change	-2.44***	5.62	-2.07***	4.48	-2.59***	5.59	
n	152		134		123		
CORE -OM	20.04	7.93	19.79	8.16	19.02	7.52	0.56
CORE-OM 4m	16.58	8.89	16.75	8.71	16.30	8.22	
Change	-3.47***	6.29	-3.03***	5.22	-2.72***	5.58	
n	153		128		120		

*** p<0.01, ** p<0.05, * p<0.1 a One Way ANOVA

Project 08/1610/154

Table 33. Multivariate linear regression : Change in health and well-being from baseline to 4 months follow-up – IAPT seen vs. IAPT not seen vs. Comparator

	Change from baseline to 4 months						
	SF-6D	**PHQ 9**	**GAD 7**	**CORE OM**	**Life satisfaction**	**Happy (frequency)**	**Happy (feelings)**
IAPT seen	0.017	0.067	0.401	-0.613	0.147	0.232**	0.067
	(0.013)	(0.656)	(0.605)	(0.689)	(0.165)	(0.100)	(0.114)
IAPT not seen	-0.010	0.767	0.576	-0.267	0.146	0.029	-0.072
	(0.013)	(0.681)	(0.625)	(0.721)	(0.172)	(0.103)	(0.119)
Baseline Score	-0.255***	-0.260***	-0.309***	-0.166***	-0.359***	-0.486***	-0.443***
	(0.040)	(0.038)	(0.043)	(0.036)	(0.038)	(0.043)	(0.044)
Age	-0.001***	0.003	-0.018	-0.003	0.007	-0.003	-0.006*
	(0.000)	(0.019)	(0.017)	(0.020)	(0.005)	(0.003)	(0.003)
Sex	0.010	0.536	-0.245	0.708	0.038	-0.088	-0.121
	(0.012)	(0.607)	(0.562)	(0.643)	(0.154)	(0.092)	(0.106)
Constant	0.226***	0.685	2.203**	0.407	1.132***	1.578***	2.010***
	(0.033)	(1.077)	(1.025)	(1.219)	(0.252)	(0.178)	(0.199)
Observations	393	409	406	399	410	412	410
R-squared	0.12	0.11	0.12	0.06	0.19	0.26	0.21

Standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Project 08/1610/154

Reference category: *Comparator site*

SF-6D, life satisfaction, happiness: positive coefficients indicate improvements over and above comparator sites

PHQ-9, GAD 7, CORE-OM: negative coefficients indicate improvements over and above comparator sites

6.3.4.3 Employment

Primary Analysis

6.2% (4/65) vs. 12% (3/25) patients had moved from unemployment into full or part time employment in Doncaster vs. Wakefield and Barnsley while the corresponding proportions were 6.7% (1/15) vs. 20% (1/5) in Newham vs. City and Hackney. Changes in unemployment were not statistically significant within the sites.

Mean (sd) days lost from work at 4 months for those who were in employment were 10.9 (22.4) compared to 14.7 (25.5) at baseline in Doncaster which was a significant reduction (Wilcoxon Signed Rank test $z = -2.75$, $p = 0.006$, $n = 114$). In Wakefield and Barnsley, this was 6.4 (15.0) vs. 8.1 (19.4) (n=46) which was not a statistically significant reduction. There were no significant differences between changes in lost employment days across the two sites. Mean days lost from work at 4 months in Newham was 4.6 (7.5) compared to 13.6 (32.2) (n= 17) at baseline which was not a statistically significant reduction. In City and Hackney, this was 7.3 (9.8) compared to 7.4 (16.9) but this lack of change was not significantly different from changes in Newham.

The proportion of individuals reporting that their emotional health affected their regular activities or work most/all of the time fell from 47% (114/245) to 28% (69/245) in Doncaster compared to 38% (38/99) to 29% (29/99) in Wakefield and Barnsley with statistically significant changes in both sites (p<0.01). In Newham, the fall was small from 40% (18/45) to 38% (17/45) which was not a statistically significant change whereas in City and Hackney this was 42% (11/26) to 12% (3/26) which was statistically significant (p<0.1).

Secondary Analysis

4.8% (2/42) vs. 7.9% (3/38) vs. 13.3% (4/30) patients had moved from unemployment into full or part time employment for patients who were seen in IAPT, IAPT not seen patients and in the comparator sites but these changes were not statistically significant within the sites.

Mean (sd) days lost from work at 4 months were 12.4 (25.1) compared to 18.3 (12.4) at baseline for IAPT seen patients which was a significant reduction (Wilcoxon Signed Rank $z = -1.92$, $p = 0.055$, $n = 71$) while for IAPT not seen patients it was 7.2 (15.0) vs. 10.1 (18.1) (Wilcoxon Signed Rank $Z= -1.86$, $p = 0.063$, $n =60$) indicating that both those who received IAPT and those who did not in IAPT sites had similar outcomes. Mean days lost from work in the comparator sites was 6.6 (14.0) at 4 months compared to 8.0 (18.8) at baseline which was not a statistically significant

reduction. There were no statistically significant differences when changes in lost employment days were compared across the sites.

Reported emotional health effects on work and regular activities most/all of the time changed from 44% (69/156) to 29% (45/156) in the IAPT seen patients compared to 47% (63/134) to 31% (41/134) in IAPT not seen and 39% (49/125) to 26% (32/125) in the comparator sites. Changes within the three sites were statistically significant (p<0.01).

6.3.4.4 Resource use

Primary analysis

As would be expected, the mean (sd) NHS psychotherapist/ counsellor use increased at 4 months in Doncaster 1.20 (2.45) compared to 0.43 (1.92) at baseline (Table 34). This increase was significantly different to changes in Wakefield and Barnsley (p <0.05). In Newham, mean (sd) NHS psychotherapist increased from 0.58 (2.45) to 1.89 (1.31) but this was matched by an increase in City and Hackney of 0.23 (1.18) to 1.73 (3.67) which meant there were no significant differences between increases in the two sites. Usage did not change significantly for any other mental health services at the IAPT sites and their matched comparator sites (Table 34).

There were statistically significant reductions in the use of GP surgery consultations in Doncaster with the mean (sd) decreasing from 3.74 (6.51) to 2.38 (3.63) (Table 35). This represented a statistically significant reduction in GP use when compared to Wakefield and Barnsley. There were no similar reductions in Newham.

There were also statistically significant reductions in the use of Accident and Emergency services in Doncaster and Newham but these were not significantly different from changes in A and E use in the matched comparator sites (Table 35).

Other unspecified health care service use had statistically significant reductions in both Doncaster (Mean difference =-0.40 sd = 2.04) and Newham (Mean difference =-0.59 sd = 2.56) with some evidence that the changes in Doncaster were significantly different from those in Wakefield and Barnsley (p=0.1) (Table 35). Changes in health visitor and social worker contacts were different between Doncaster and its matched sites and although these differences were small, they were statistically significant (Table 35).

Secondary analysis

As would be expected, NHS psychotherapist/counsellor use increases were mainly for IAPT seen patients (Mean difference =1.40 sd = 3.56) although there was some increase in NHS psychotherapists use in the IAPT not-seen patients (mean difference =0.24, sd = 3.07) and respondents from the comparator sites (mean difference =0.24, sd = 2.81) (Table 36). There were no other statistically differences in mental health service use across the three groups.

GP surgery consultations had statistically significant reductions in both IAPT seen patients (mean difference =-0.92, sd = 3.88) and not-seen patients (mean difference = -1.51, sd = 6.53) but there were no significant differences across these 2 groups and the comparator sites group (Table 37).

Accident and Emergency reductions were statistically significant in both IAPT seen patients (mean difference =-0.10, sd = 0.59) and IAPT not seen patients (mean difference =-0.31, sd = 2.62) groups with no statistically significant differences across the 3 groups. This maybe an indication that these reductions may have been as a result of changes across the site as opposed to specific IAPT related gains.

Table 34. Change in mental health service use 4m: IAPT and comparator matched sites

No of times seen	Doncaster		Wakefield & Barnsley		MW Test	Newham		City & Hackney		MW Test
	mean	sd	mean	sd	P Value	mean	sd	mean	sd	P Value
NHS Psychotherapist or counsellor	0.43	2.25	0.45	2.05	0.01	0.58	2.45	0.23	1.18	0.04
NHS Psychotherapist or counsellor 4m	1.20	2.45	0.36	1.18		1.89	4.34	1.73	3.67	
	0.78***	3.00	-0.09	2.39		1.31**	5.02	1.50**	3.84	
Psychiatrist	0.09	0.76	0.07	0.33	0.13	0.38	1.09	0.23	0.86	0.69
Psychiatrist 4m	0.06	0.36	0.15	0.61		0.31	0.63	0.69	3.15	
	-0.03	0.80	0.08	0.65		-0.07	1.03	0.46	3.26	
Community Psychiatric Nurse	0.45	6.15	0.12	0.67	0.10	0.16	0.77	0.04	0.20	0.62
Community Psychiatric Nurse 4m	0.20	2.44	0.38	1.90		0.20	0.73	0.00	0.00	
	-0.25	3.74	0.26	2.02		0.04	0.52	-0.04	0.20	
Psychologist	0.11	1.12	0.00	0.00	0.71	0.27	1.37	0.04	0.20	1.00
Psychologist 4m	0.05	0.34	0.00	0.00		0.20	0.84	0.04	0.20	
	-0.07	1.17				-0.07	1.64	0.00	0.28	
Other Mental health service	0.26	1.10	0.03	0.17	0.59	0.41	1.85	0.00	0.00	0.10
Other Mental health service 4m	0.34	1.39	0.25	1.47		0.02	0.15	0.04	0.20	
	0.08	1.67	0.22	1.48		-0.39*	1.86	0.04	0.20	
Private Psychotherapist or counsellor	0.46	2.47	0.25	1.24	0.26	0.16	0.74	0.50	1.82	0.23
Private Psychotherapist or counsellor 4m	0.55	2.50	0.10	0.59		1.00	3.21	0.35	1.20	
	0.09	3.33	-0.15	1.21		0.84	2.99	-0.15	2.01	

No of times seen	Doncaster		Wakefield & Barnsley		MW Test	Newham		City & Hackney		MW Test
	mean	sd	mean	sd	P Value	mean	sd	mean	sd	P Value
Voluntary sector mental health	0.20	1.59	0.31	2.13	0.51	0.42	2.42	0.31	1.57	0.15
Voluntary sector mental health 4m	0.36	2.88	0.40	2.39		0.38	2.39	0.46	2.35	
	0.16	3.30	0.09	2.41		-0.04	0.30	0.15	0.78	

*** p<0.01, ** p<0.05, * p<0.1 Wilcoxon Signed Rank Test

MW – Mann Whitney test

Table 35. Change in other health service use 4m: IAPT and comparator matched sites

	Doncaster		Wakefield & Barnsley		MW test	Newham		City & Hackney		MW Test
	mean	sd	mean	sd	P value	mean	sd	mean	sd	P value
GP	3.74	6.51	2.60	3.34	0.02	2.36	3.50	2.81	3.11	0.79
GP 4m	2.38	3.63	2.20	2.63		2.02	2.54	2.92	4.10	
Change	-1.36***	5.60	-0.40	3.24		-0.33	2.88	0.12	3.99	
GP phone advice	0.40	1.20	0.29	1.17	0.72	0.24	0.80	0.46	1.10	0.32
GP phone advice 4m	0.29	0.82	0.20	0.74		0.33	0.83	0.35	1.02	
Change	-0.11	1.32	-0.09	1.35		0.09	1.02	-0.12	1.34	
GP home visits	0.04	0.23	0.00	0.00	0.52	0.00	0.00	0.00	0.00	1.00
GP home visits 4m	0.04	0.36	0.02	0.20		0.00	0.00	0.00	0.00	
Change	0.00	0.38	0.02	0.20						
GP practice nurse	0.62	1.56	0.52	1.14	0.55	0.56	1.83	1.50	6.24	0.95
GP practice nurse 4m	0.58	1.20	0.54	1.22		0.22	0.60	1.00	3.20	
Change	-0.04	1.96	0.02	1.53		-0.33	1.85	-0.50	3.34	
NHS Direct	0.27	1.06	0.19	0.60	0.44	0.40	1.05	0.12	0.43	0.68
NHS Direct 4m	0.12	0.45	0.18	0.67		0.38	0.94	0.19	0.63	
Change	-0.15	1.12	-0.01	0.67		-0.02	1.18	0.08	0.63	
A and E	0.29	1.97	0.22	0.64	0.93	0.27	0.54	0.12	0.59	0.14
A and E 4m	0.09	0.31	0.14	0.59		0.09	0.36	0.08	0.27	
Change	-0.20**	1.99	-0.08	0.61		-0.18**	0.53	-0.04	0.45	

	Doncaster		Wakefield & Barnsley		MW test	Newham		City & Hackney		MW Test
	mean	sd	mean	sd	P value	mean	sd	mean	sd	P value
Day hospital	0.18	1.18	0.08	0.37	0.74	0.07	0.25	0.04	0.20	0.88
Day hospital 4m	0.20	1.40	0.13	0.46		0.13	0.46	0.08	0.27	
Change	0.02	1.83	0.05	0.58		0.07	0.50	0.04	0.34	
Health visitor	0.11	0.69	0.06	0.31	0.08	0.13	0.66	0.00	0.00	0.66
Health visitor 4m	0.05	0.44	0.11	0.51		0.13	0.63	0.00	0.00	
Change	-0.06	0.72	0.05	0.46		0.00	0.37			
Outpatient attendances	0.42	2.16	0.38	1.68	0.17	0.64	2.12	0.19	0.80	0.19
Outpatient attendances 4m	0.62	2.01	0.39	1.62		0.20	1.20	0.23	0.71	
Change	0.20**	2.48	0.01	2.05		-0.44	2.48	0.04	1.11	
Social Worker	0.06	0.35	0.02	0.20	0.06	0.13	0.63	0.00	0.00	0.65
Social Worker 4m	0.05	0.44	0.12	0.67		0.27	1.50	0.00	0.00	
Change	-0.01	0.38	0.10	0.70		0.13	1.10			
Other health service	0.46	2.06	0.04	0.32	0.10	0.73	2.74	0.04	0.20	0.41
Other health service 4m	0.06	0.37	0.22	1.62		0.14	0.90	0.00	0.00	
Change	-0.40***	2.04	0.18	1.67		-0.59*	2.56	-0.04	0.20	

*** p<0.01, ** p<0.05, * p<0.1 based on Wilcoxon Signed Rank Test

MW – Mann Whitney test

Table 36. Change in mental health service use 4m: IAPT seen vs. IAPT not seen vs. Comparator

No of times seen	IAPT seen		IAPT site not seen		Comparator site		Kruskal Wallis Test
	mean	sd	mean	sd	mean	sd	P value
NHS Psychotherapist or counsellor	0.39	1.68	0.52	2.33	0.40	1.90	0.00
NHS Psychotherapist or counsellor 4m	1.79	3.35	0.76	1.93	0.64	2.03	
	1.40***	3.56	0.24**	3.07	0.24*	2.81	
Psychiatrist	0.08	0.58	0.19	1.03	0.10	0.49	0.54
Psychiatrist 4m	0.09	0.45	0.11	0.40	0.26	1.52	
	0.01	0.65	-0.08	1.01	0.16	1.58	
Community Psychiatric Nurse	0.11	0.62	0.75	8.27	0.10	0.60	0.38
Community Psychiatric Nurse 4m	0.08	0.43	0.34	3.28	0.30	1.70	
	-0.03	0.57	-0.41	5.01	0.20	1.81	
Psychologist	0.19	1.48	0.08	0.64	0.01	0.09	0.81
Psychologist 4m	0.06	0.40	0.09	0.51	0.01	0.09	
	-0.13	1.53	0.01	0.82	0.00	0.13	
Other Mental health service	0.36	1.48	0.19	0.89	0.02	0.16	0.25
Other Mental health service 4m	0.31	1.49	0.28	0.10	0.20	1.31	
	-0.05	2.12	0.08	1.05	0.18	1.32	
Private Psychotherapist or counsellor	0.49	2.88	0.33	1.35	0.30	1.38	0.25
Private Psychotherapist or counsellor 4m	0.75	3.09	0.46	1.95	0.15	0.76	
	0.26	4.24	0.13	1.64	-0.15	1.40	

Voluntary sector mental health	0.33	2.16	0.12	1.07	0.31	2.02	0.97
Voluntary sector mental health 4m	0.18	1.35	0.56	3.85	0.41	2.37	
	-0.15	1.79	0.44	4.01	0.10	2.17	

Kruskal Wallis Test of change in health service use across the 3 groups;
*** p<0.01, ** p<0.05, * p<0.1 - Wilcoxon Signed Rank Test Paired differences between baseline and 4 months in each site.

Table 37. Change in other health service use 4m: IAPT seen vs. IAPT not seen vs. comparator

	IAPT seen		IAPT site not seen		Comparator site		Kruskal Wallis Test
	mean	sd	mean	sd	mean	sd	P value
GP Surgery Consultation	3.20	3.31	3.89	8.30	2.64	3.28	0.15
GP Surgery Consultation 4m	2.28	3.02	2.38	3.96	2.35	2.99	
	-0.92***	3.88	-1.51***	6.53	-0.29	3.40	
GP phone advice	0.35	1.17	0.39	1.13	0.33	1.15	0.67
GP phone advice 4m	0.26	0.81	0.33	0.83	0.23	0.80	
	-0.09	1.27	-0.06	1.30	-0.10	1.34	
GP home visits	0.02	0.18	0.04	0.24	0.00	0.00	0.49
GP home visits 4m	0.00	0.00	0.07	0.48	0.02	0.18	
	-0.02	0.18	0.03	0.47	0.02	0.18	
GP practice nurse	0.68	1.81	0.52	1.33	0.72	3.00	0.69
GP practice nurse 4m	0.46	0.98	0.59	1.29	0.63	1.80	
	-0.22	1.99	0.07	1.88	-0.09	2.03	
NHS Direct	0.30	1.09	0.27	1.03	0.17	0.57	0.47
NHS Direct 4m	0.14	0.54	0.19	0.59	0.18	0.66	
	-0.16*	1.13	-0.09	1.14	0.01	0.66	
A and E	0.19	0.55	0.39	2.60	0.20	0.63	0.62
A and E 4m	0.10	0.34	0.08	0.30	0.13	0.54	
	-0.10**	0.59	-0.31**	2.62	-0.07	0.58	

	IAPT seen		IAPT site not seen		Comparator site		Kruskal Wallis Test
	mean	sd	mean	sd	mean	sd	P value
Day hospital	0.17	1.20	0.16	0.95	0.07	0.34	0.81
Day hospital 4m	0.14	0.46	0.25	1.84	0.12	0.43	
	-0.03	1.26	0.10	2.08	0.05	0.53	
Health visitor	0.10	0.58	0.13	0.79	0.05	0.28	0.34
Health visitor 4m	0.05	0.35	0.09	0.59	0.09	0.46	
	-0.05	0.49	-0.04	0.84	0.04	0.41	
Outpatient attendances	0.58	2.70	0.32	1.24	0.34	1.54	0.59
Outpatient attendances 4m	0.53	1.52	0.59	2.29	0.36	1.48	
	-0.05	2.81	0.27	2.04	0.02	1.89	
Social Worker	0.05	0.32	0.09	0.48	0.02	0.18	0.05
Social Worker 4m	0.13	0.94	0.03	0.27	0.10	0.60	
	0.08	0.68	-0.06	0.36	0.08	0.63	
Other health service	0.76	2.85	0.21	0.84	0.04	0.30	0.05
Other health service 4m	0.08	0.57	0.06	0.39	0.18	1.44	
	-0.68***	2.77	-0.15*	0.92	0.13	1.48	

Kruskal Wallis Test of change in health service use across the 3 groups.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$ - Wilcoxon Signed Rank Test Paired differences between baseline and 4 months in each site.

Project 08/1610/154

6.1.14   Follow-up at 8 months

Changes were similar between sites for health, well-being and resource use in the primary and secondary comparisons at 8 months. The detailed results are not presented in the main report, but can be found in the Appendix (Appendix 6 Table 6.12 to 6.19).

Health and well-being continued to improve modestly in all groups (Appendix 6 Table 6.12 to 6.15) but there were no significant differences between matched IAPT and comparator sites although IAPT seen patients have slightly larger happiness levels compared to the comparator sites even after controlling for baseline values, age and gender (Appendix 6 Table 6.15).

There were non-significant increases in NHS psychotherapy use for the two demonstration sites and Wakefield and Barnsley but there was a significant increase in City and Hackney ($p<0.1$) which is unsurprising given that this site became a Wave 1 IAPT site (Appendix 6 Table 6.16). There was also a significant reduction in GP surgery visits when Doncaster was compared to its matched site ($p<0.05$) (Appendix 6 Table 6.17) . However, the other resource use did not change significantly.

6.1.15   Examining potential bias in IAPT seen cohort patients

6.3.6.1 Patient Reported Outcome Measures (PROMS)

There were 190 cohort patients who were seen in IAPT services in Doncaster (n = 142) and Newham (n= 48) (Appendix 6 Table 6.20). Comparisons of the IAPT service scores (2009) for cohort patient vs. non-cohort patients indicated that there were no substantial differences between patients who were recruited into the cohort study and the rest of the patients who were seen in IAPT services (Appendix 6 Table 6.21 and 6.22).

However, the cohort study scores at the end of 4 months and 8 months were different from those reported in Chapter 5 where IAPT service intention to treat (ITT) scores were used (Table 38). Baseline scores were similar when all the IAPT patients scores were compared to the two the IAPT seen cohort patient scores ($16.1_{(ITT)}$ vs. $16.2_{(cohort\ 1)}$ vs. $15.7_{(cohort\ 2)}$)[48] but follow-up scores were higher for the cohort study scores, particularly at 4 months, ($9.7_{(ITT)}$ vs. $13.0_{(cohort\ 1)}$ vs. $11.8_{(cohort\ 2)}$) which seemed to indicate that cohort patients had poorer outcomes. There was therefore concern that the cohort study scores were biased in some way compared to the IAPT

---

[48] Cohort 1 refers to the sample at 4 months; Cohort 2 refers to the sample at 8 months

service scores. This discrepancy could be explained by: 1) the inclusion of patients who were not seen in IAPT services in the cohort sample; 2) differences in the timing of assessments between the cohort survey, that was at 4 and 8 months after baseline and the IAPT service, that that was determined by treatment schedules; 3) place of administration, since IAPT service assessments were undertaken on site and the cohort assessments were completed in the patients own home; or 4) selection bias – where those with a successful outcome were less likely to respond to the cohort survey.

Exclusion of patients not seen in IAPT

Focusing on patients who were seen in the service did not make a difference to the 8 month follow-up period. Cohort study scores at the 8 month follow up period for patients who were seen in the service were still considerably higher than the IAPT intention to treat scores ($9.7_{(ITT)}$ vs. $11.9_{(cohort\ 2)}$).

Taking into account timing issues

However, the IAPT intention to treat scores were based on the April 2009 download data and this excludes measures for the cohort study taken after this point. The April 2010 download data indicates that IAPT service scores for the cohort patients are similar to those of the rest of the IAPT service patients at follow-up ($9.7_{(ITT)}$ vs. $10.4_{(IAPT\ 2010)}$) but this was still lower than the cohort study scores of 11.9.

Another timing issue was the fact that the IAPT service scores were not necessarily taken at the same time as the study period for cohort patients. Some patients had been seen in the service prior to their entry into the cohort study. Last scores from the IAPT service were for the last recoded session which may have occurred at a point outside the 8 month follow-up. Using first and last scores from the IAPT service (2009) that were closest to the cohort patients baseline and 8 month follow-up dates, lowered both the baseline and follow-up scores. There were significant differences between the cohort baseline and the IAPT April 2009 scores that occurred within the study period ($15.6_{(cohort\ 2)}$ vs. $14.5_{(IAPT\ study\ period)}$; $t_{(91)}$ 2.43 $p<0.05$). Follow up scores were also different ($10.4_{(cohort\ 2)}$ vs. $9.4_{(IAPT\ study\ period)}$) but these differences were not statistically significant. Timing differences may therefore have contributed to the differences in overall IAPT scores and cohort study scores.

Focusing on those seen in IAPT and aligning the timing of assessments more accurately resulted in a smaller difference in follow-up scores. The remaining difference is 1.5 on the PHQ 9 and this could have been accounted for by the place of administration or selection bias. This represented a potential downward bias of around 20% on the change in

self-reported scores (1/5/6.4) in the cohort study, although it is not possible to say what proportion of this was attributable to the IAPT intervention.

Table 38. PHQ 9 scores for patients in Doncaster IAPT

		Baseline	Follow-up	Change	N
Benchmarking chapter – all IAPT service patients					
PHQ-9	Intention to treat	16.1 (6.2)	9.7 (7.5)	6.4	4616
Cohort Patients (all)					
PHQ 9	Cohort study scores (4 months)	16.2 (7.1)	13.0 (7.7)	3.2	244
	Cohort study scores (8 months)	15.7 (7.1)	11.8 (7.6)	3.9	231
Cohort Patients (IAPT seen)					
PHQ 9	Cohort 2 study scores (8 months)	15.9 (7.1)	11.9 (7.6)	4.0	124
	IAPT April 2010 scores	16.4 (6.0)	10.4 (7.3)	6.0	103
	IAPT April 2009 (study period) scores	14.8 (6.7)	9.4 (6.7)	5.3	72

6.3.6.2 Resource Use

Download data from the IAPT services indicated that of the IAPT seen patients, 29 had only one contact (6 in Newham and 23 in Doncaster) during the study period. The mean (sd) number of contacts for the IAPT seen patients was 4.0 (3.0) with a range from 0 to 12 and 10.9 (8.2) with a range of 0 to 34 in Doncaster and Newham respectively. Self-reported contacts from the study questionnaire were 2.3 (3.4) (range 0 to 15) and 3.6 (6.7) (range 0 to 23) for IAPT seen patients in Doncaster and Newham over the entire study period which indicates considerable under-reporting of service use especially in Newham. The mean differences in number of contacts that were self-reported and those from the IAPT services were statistically significant in both Doncaster ($t_{(92) = 8.24}$, p = 0.000) and Newham ($t_{(26)}$, p = 0.000).

6.1.16   Costing of IAPT services

6.3.7.1 Cost per year

The PCTs of the IAPT services provided financial statements for the years 2007-8 and 2008-9 covering staff, training, equipment and overheads (Appendix 6 Tables 6.23 to 6.26). The total costs of the IAPT services to the NHS in Doncaster in those years were £1,652,525 (£1,723,619 in 2008/09

prices) and £1,795,363 respectively and in Newham were £1,262,953 (£1,317,287 in 2008/09 prices) and £1,313,704.

6.3.7.2 Total patient and session numbers

These cost data were combined with the total activity data on patients who were seen in both services from the service download data for each financial accounting year (which runs from 6th April to 5th April) so that the unit costs for the IAPT service could be estimated (Appendix 6 Table 6.27 and 6.28 Doncaster ; Appendix 6 Tables 6.29-6.31 Newham).

The number of patients referred to Doncaster over this period was 3920 and 4120 in 2007/08 and 2008/09 and the equivalent numbers were 1153 and 1628 in Newham. The number of patients who attended at least one session in Doncaster over this period was 3,410 [49] and 3,021 whereas in Newham it was 1,078 and 1,713

Total patient attributed time was 430,650 and 384,045 minutes in Doncaster and 247,618 and 544,330 in Newham over those time periods. Total time in Newham was based on the sum of all time recorded (clinical, administration and supervision). Clinical time alone in Newham was 185,973 and 447,930 minutes respectively. Based on the total time in Doncaster and clinical time in Newham, the average cost per minute over the two years was £4.33 in Doncaster and £5.01 in Newham. These average costs were used to estimate the cost of the IAPT service in the next section.

6.3.7.3 Costs of all services

Primary analysis

Mean total NHS and PSS costs per patient were £1,185 in Doncaster and £1,507 in Newham weighted over activity in the two years (Table 39 and 40)[50]. This was more expensive than the costs associated with treatment in their matched sites where costs were £924 and £946. Mental health costs[51] in Doncaster were £345 compared to £864 in Newham compared to £80 and £161 in Wakefield & Barnsley and City & Hackney. The larger costs in

---

[49] The information from 2007/08 for Doncaster had 73.7% that did not have session outcome data. These appear to be earlier in the year before changes in the IAPT IT system were implemented to allow outcomes to be recorded more accurately. This number was therefore based on those sessions that had time recorded and was therefore an overestimate of attendances .

[50] Baseline average costs are reported in Appendix 6 Table 33

[51] Mental health costs include: NHS physiotherapist or counsellor/ IAPT service use, psychiatrist use, community psychiatric nurse, and psychologist.

Newham are a reflection of the high intensity service that was offered there particularly in the first year where fewer individuals were seen with a higher number of sessions per individual.

Secondary Analysis

Mean total NHS and PSS costs for per patient for those who were seen in IAPT in either Doncaster or Newham were £1,451 compared to £979 and £929 for not-seen IAPT patients and comparator sites patients (Table 41). £680 of this cost was mental health compared to £126 and £98 for the two other groups. This suggest a marginal cost attributable to IAPT of £ 519.

6.3.7.4 Costs of lost employment

The cost of lost employment was higher in Doncaster than its matched site at £667 compared to £388 (Table 39).The opposite was true in Newham where lost employment was £151 compared to £363 in City and Hackney (Table 40). The large differences in lost employment days were a result of skewed data as trimmed means (5%) for Doncaster, Wakefield and Barnsley, Newham and City and Hackney were £278, £211, £91 and £256 indicating that the data were highly skewed by outliers.

Lost employment costs were highest for those who were seen in IAPT at £695 compared to £455 and £383 for those who did not receive IAPT interventions (Table 41). 5% trimmed means were £284, £180, and £219 for the IAPT seen, not seen and comparator sites.

6.1.17   Cost effectiveness

6.3.8.1 Primary analysis

Cost effectiveness was examined using individual patient level SF-6D and cost data. The mean cost differences of patients in the Doncaster and Newham cohorts using IAPT costs were £263 (95% CI: -£258 to 779) and £561 (95% CI: -£333 to 1,454) respectively. These were associated with QALY differences of 0.007 (95% CI: -0.006 to 0.021) and -0.002 (95% CI: -0.035 to 0.031) respectively compared to their control sites samples. This resulted in an incremental cost effectiveness ratio (ICER) of £37,571 per QALY when Doncaster was compared to Wakefield and Barnsley. Newham QALY outcomes were dominated by City and Hackney and so an ICER was not calculated.

All estimates were associated with very large degrees of uncertainty that was reflected in the wide 95% confidence intervals that included positive and negative values. The cost effectiveness acceptability curves (Figure 14) indicate that the probability that IAPT was cost-effective was below 40% at a cost of £30,000 per QALY.

6.3.8.2 Secondary Analysis

The secondary comparison of IAPT seen patients to comparator cases resulted in an incremental cost of £519 (95% CI: £20 to 1,025), which was statistically significant at the 5% level. There was a difference in QALY gain of 0.013 (95% CI: -0.002 to 0.029), which was nearly significant (Table 41). These values resulted in a cost per QALY ratio of £39,923, but this was associated with a very high degree of uncertainty. The probability that IAPT was cost effective when comparing those who received the intervention and the comparator site was around 8% at £20,000 per QALY and around 38% at £30,000 per QALY (Figure 15).

6.3.8.3 Sensitivity analysis

EQ-5D

Sensitivity analysis using predicted EQ-5D scores indicated that there was a QALY gain of 0.013 (95% CI: -0.007 to 0.033) when Doncaster was compared to Wakefield and Barnsley, which was nearly significant (Table 39). This resulted in a cost per QALY ratio of £20,230. The probability that this IAPT service was cost effective was 43% for a willingness to pay for a QALY of £20,000 and 57% at £30,000 (Figure 16). QALY gains for Newham were still less than those for City and Hackney when predicted EQ-5D scores were used (Table 40).

When the IAPT seen group was compared to the comparator group, EQ-5D QALY gains were 0.020 (95% CI: -0.0002 to 0.041) (Table 41) which led to a cost per QALY ratio of £25,950. The probability that IAPT was cost effective when focusing on the IAPT seen patients was 58% at £30,000 per QALY (Figure 17).

National unit costs for IAPT

When national unit costs were used in place of IAPT costs, the mean total NHS and PSS cost per patient were £1,042 (95% CI: £749 to 1,334) in Doncaster and £1,176 (95% CI: £729 to 1,523) in Newham. This represented an incremental cost of £117 (95% CI: -£396 to 631) for the IAPT service in Doncaster resulting in an ICER of £16,714 per QALY using SF-6D to calculate QALYs.

The IAPT seen group mean total NHS and PSS costs when national unit costs were used were £1,133 (95% CI: £875 to 1,392) which represented an incremental cost of £ 204 (95% CI: -£346 to 755) over the comparator site. This resulted in an ICER of £15,692 per QALY.

Table 39. Costs and QALYs at 4 and 8 months : Doncaster vs. Wakefield & Barnsley

	Doncaster		Wakefield & Barnsley		Difference	
	Mean	**95% CI**	**Mean**	**95% CI**		**95% CI**
Total NHS and PSS costs	1,185	888 to 1,481	924	566 to 1,282	263	-258 to 779
Total lost employment costs	667	395 to 940	388	176 to 600	279	-65 to 624
Mental health services costs	345	282 to 407	80	36 to 124	265	189 to 341
SF-6D QALY Gain/Loss	0.025	0.018 to 0.033	0.018	0.007 to 0.029	0.007	-0.006 to 0.021
EQ-5D QALY Gain/Loss	0.037	0.026 to 0.048	0.024	0.009 to 0.039	0.013	-0.007 to 0.033
n	212		85			

Table 40. Costs and QALYs at 4 and 8 months: Newham vs. City & Hackney

	Newham		City & Hackney		Difference	
	Mean	**95% CI**	**Mean**	**95% CI**		**95% CI**
Total NHS and PSS costs	1,507	904 to 2109	946	339 to 1,553	561	-333 to 1,454
Total lost employment costs	151	40 to 262	363	70 to 656	-212	-522 to 98
Mental health services costs	864	433 to 1295	161	59 to 264	703	261 to 1,114
SF-6D QALY Gain/Loss	0.021	0.001 to 0.043	0.023	0.001 to 0.043	- 0.002	-0.035 to 0.031
EQ-5D QALY Gain/Loss	0.028	-0.003 to 0.059	0.035	0.001 to 0.069	-0.007	-0.054 to 0.040
n	40		24			

Table 41. Costs and QALYs at 4 and 8 months : IAPT seen vs. IAPT not-seen vs. Comparator

	IAPT seen		IAPT not seen		Comparator		Difference (IAPT seen vs. Comparator)	
	Mean	95% CI	Mean	95% CI	Mean	95% CI	Mean	95% CI
Total NHS and PSS costs	1,451	1,164 to 1,739	979	507 to 1,451	929	624 to 1,234	519	20 to 1,025
Total lost employment costs	695	338 to 1052	455	176 to 734	383	207 to 558	313	-84 to 709
Mental health services costs	680	537 to 833	126	76 to 177	98	57 to 139	582	433 to 730
SF-6D QALY Gain/loss	0.032	0.023 to 0.042	0.015	0.005 to 0.026	0.019	0.010 to 0.029	0.013	-0.002 to 0.029
EQ-5D QALY Gain/loss	0.047	0.036 to 0.064	0.023	0.007 to 0.039	0.026	0.013 to 0.040	0.020	-0.0002 to 0.041
n		137		115		109		

Figure 14. **SF-6D Cost Effectiveness Acceptability Curve: Doncaster vs. Wakefield & Barnsley**

Cost effectiveness acceptability curve
Doncaster vs. Wakefield & Barnsley
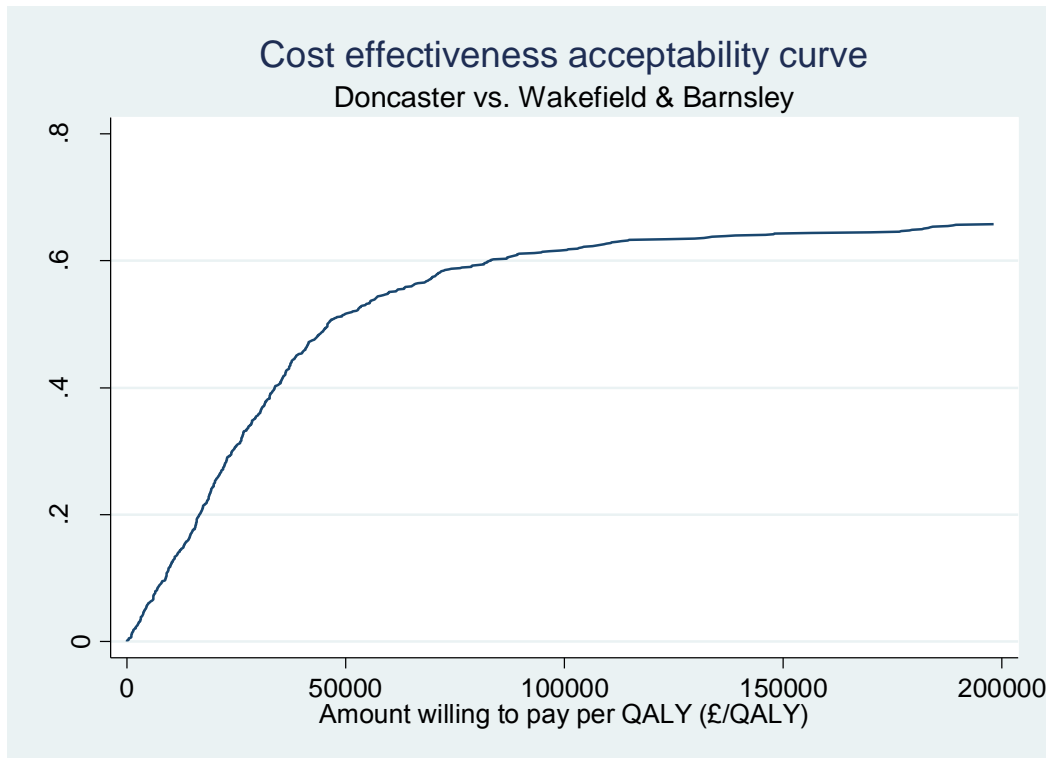
Figure 15.   SF-6D Cost Effectiveness Acceptability Curve: IAPT seen vs. Comparator

Cost effectiveness acceptability curve
IAPT seen vs Comparator
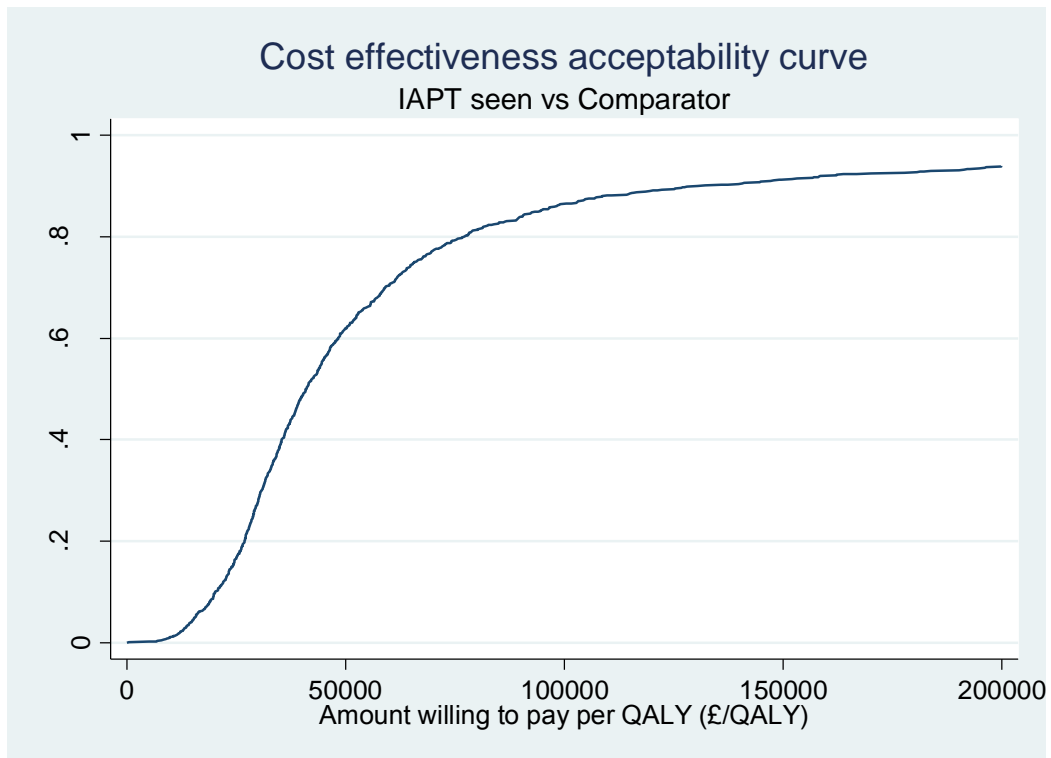
Figure 16.   EQ-5D Cost Effectiveness Acceptability Curve: Doncaster vs. Wakefield & Barnsley

Cost effectiveness acceptability curve
Doncaster vs. Wakefield & Barnsley
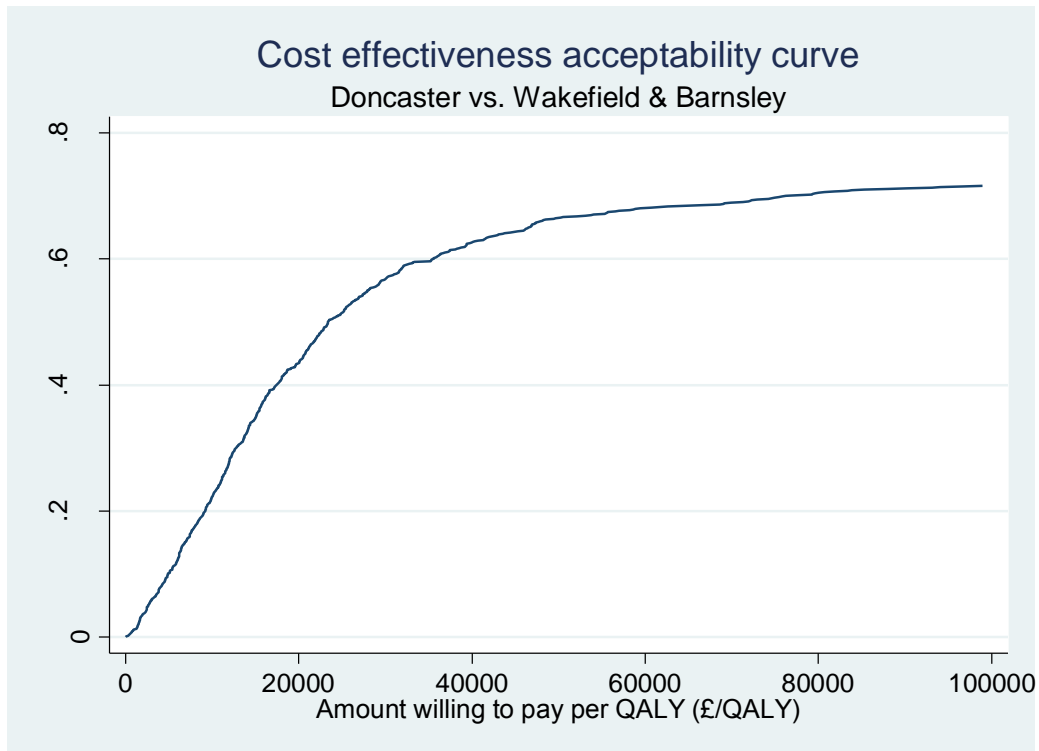
Figure 17. **EQ-5D Cost Effectiveness Acceptability Curve: IAPT seen vs.**
Comparator
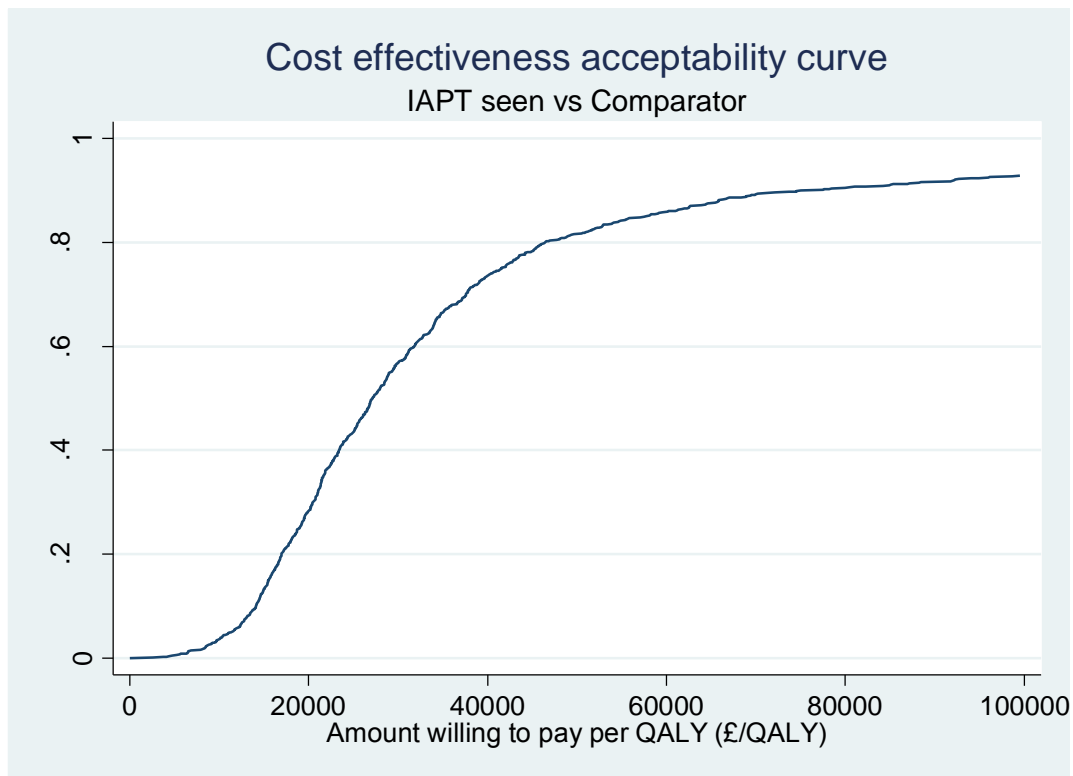
Discussion

This study sought to test the hypothesis that the IAPT services in Doncaster and Newham were a cost effective use of NHS resources. To answer this question, it took an innovative approach that did not rely on using routinely collected data on those patients who used these services. Over a 2 year period it recruited patients from selected practices from the IAPT sites and matched controlled sites. This was aimed at avoiding problems with selection bias and lack of control from a routine data set. It also provided patients' views of their health away from the setting of the IAPT sites.

This comparative cohort study found quite small differences between the improvements in patients recruited in general practice in the IAPT sites and those from the matched comparator PCTs. The differences tended to favour Doncaster over the controls, for example compared to its matched controls the difference in SF-6D was 0.015 at 4 months and 0.01 after controlling for background variables. None of these differences were significant at 4 months and by 8 months had disappeared. In an attempt to focus more on

those most likely to benefit from IAPT, we undertook a secondary analysis of those cases actually seen in IAPT compared to those recruited from IAPT practices and not seen at the IAPT sites and the control. The differences in the SF-6D were 0.027 and 0.002 at 4 and 8 months respectively. These differences were not significant, except for happiness frequency that had a significantly better score in those seen in IAPT. These differences resulted in small QALY differences between the IAPT sites and their controls, even after focusing on those seen in IAPT. The confidence intervals lay either side of zero, though the QALY gain for those seen in IAPT was nearing significance at the 5% level.

Access to IAPT services led to significant increases in the use of an NHS Psychotherapist or counsellor and reductions in GP use as well as other unspecified health services in Doncaster that were significantly different from changes in equivalent service use in Wakefield and Barnsley. NHS psychotherapy/counselling use also increased significantly in both Newham and City and Hackney (where IAPT services were introduced during the study) and GP service use fell slightly in Newham but this was not a significant decrease.

Service costs were £263 (95% CI: -£258 to £779) higher in Doncaster compared to Wakefield and Barnsley and £561 (95% CI: -£333 to £1,451) higher in Newham compared to City and Hackney over 8 months for IAPT. These additional costs of IAPT generated 0.007 (95% CI: -0.006 to 0.021) additional QALYS in Doncaster but were associated with QALY losses in Newham, -0.002 (95% CI: -0.035 to 0.031), compared to their respective matched sites. This resulted in an incremental cost effectiveness ratio (ICER) of £37,571 per QALY when Doncaster was compared to Wakefield and Barnsley but the probability that IAPT was cost effective at £30,000 per QALY was below 40%. Focusing on patients seen in IAPT gave an ICER value of £39,923 which was based on an incremental cost of £519 (95% CI: £20 to 1,025), and a QALY gain of 0.013 (95% CI: -0.002 to 0.029), with again a low probability of cost effectiveness.

However, the results were sensitive to the method for valuing health states and the use of national unit costs data rather than local costs. Using EQ-5D values predicted from the SF-6D based on a regression model estimated on another data set led to ICER values of £20,230 and £25,950 per QALY for Doncaster and IAPT seen patients which was within the NICE threshold. Using national costs rather than IAPT costs resulted in values of £16,714 and £15,692 per QALY for Doncaster and IAPT seen patients which was also within the NICE threshold. There is a reasonable argument for using EQ-5D values since this makes them more consistent with the NICE threshold. Whether national unit cost data should be applied to the IAPT service is more debateable. Local costs may reflect some 'learning effects' and start

up costs, however, the IAPT service has special features that may not be reflected in the national cost data.

There were no significant changes in employment status but lost employment days fell significantly in Doncaster, Wakefield and Barnsley and City and Hackney but not in Newham. Lost employment costs changes were higher for Doncaster compared to Wakefield & Barnsley, £279 (95% CI: -£65 to £624) but lower for Newham compared to City & Hackney, -£212 (95% CI: -£522 to £98) but these lost employment cost data were highly skewed especially in Doncaster.

The innovative study design was undermined by the poor response rate amongst the patients invited to participate of 12%. This has implications for the representativeness of the samples and power. Care must be taken in generalising from this cohort to the IAPT services in Doncaster and more so in Newham. Newham and City and Hackney suffered from having a substantially smaller sample, 41 and 27 respectively, compared to 207 in Doncaster and 85 in Wakefield and Barnsley. Furthermore, the control site for Newham, City and Hackney, became a Wave 1 IAPT site during the study and this was reflected in the comparable use of psychological services observed between these sites.

Another concern about the results in the cohort sample was the differences with the results from the benchmarking analysis presented in Chapter 5. While the mean PHQ-9 baseline score in all IAPT patient in Doncaster was similar, the follow-up scores were substantially lower than in the cohort study. This might suggest that the cohort study scores underestimated the gain from IAPT. Taking into account those patients who were seen in IAPT and timing issues reduced some of the differences but there was still a difference of 1.5 in the PHQ-9 between IAPT study scores and the rest of IAPT. This indicated that improvements in PHQ 9 scores may have been underestimated by up to 20% which may also have affected the SF-6D scores and associated QALY gains.

In addition, there was evidence of under-reporting of contacts with an NHS Psychotherapist/counsellor in the cohort study. The rest of the self-reported health service use may also have been subject to some form of systematic bias and this will have an impact on associated costs. This view is supported by other studies. In their review of the use of studies of the reliability of patients self-reported health service utilisation, Evans and Crawford [79] found that under-reporting was more common than over-reporting for hospital and outpatient consultations although not all the studies they reviewed had this problem. Petrou et al [80] also found that under-reporting was a problem for community services such as general practitioners.

The results from this comparative cohort study indicate that the Doncaster IAPT demonstration site provided a service that could be cost-effective within the usual NICE threshold range of £20,000-30,000, but there was considerably uncertainty surrounding the costs and outcome differences and it was somewhat undermined by the low response rate to the patient questionnaire (though comparisons with the IAPT suggest this may have resulted in an underestimate of the cost effectiveness of this service). It is not possible to comment on the cost effectiveness of the Newham service since the numbers were too low and the comparator site adopted an IAPT service during the study.

7 General practitioner treatment and referrals

7.1 Background

One of the potential benefits of the development of the IAPT programme is a reduction in antidepressant prescribing by general practitioners, because the improved access allows patients a choice of psychological treatment instead of having to accept medication as the only available option [81-83].

In making the case for psychological treatment centres, Layard also suggested there would be significant savings to the National Health Service, because fewer people would be referred for supposed physical illnesses whose medically unexplained symptoms were due to anxiety or depression [83].

This is a reasonable suggestion since the syndrome of depression includes physical symptoms such as fatigue, insomnia, appetite loss and chronic pain. Patients whose depression is not treated effectively may be referred for investigation for possible physical causes of their persisting symptoms such as anaemia, occult cancers, joint and back problems, or neurological disorders. Similarly, anxiety symptoms overlap with symptoms of cardiovascular disease or thyrotoxicosis (palpitations, chest pains, faintness, flushing and sweating), respiratory disease (shortness of breath, hyperventilation), gastrointestinal disease (choking, feeling a lump in the throat, dry mouth, nausea, vomiting, or diarrhoea), neurological disease (dizziness, headache, parasthesiae, or vertigo), and musculoskeletal disease (muscle ache, muscle tension, tremor, or restlessness).

The symptoms of panic attacks, such as palpitations, tachycardia, shortness of breath and chest pain, may lead some individuals to think that they are experiencing a potentially life threatening event, such as a heart attack. This often results in presentation to A&E departments. It has been estimated that between 18% and 25% of patients who present to emergency or outpatient cardiology settings meet the criteria for panic disorder [84].

In this chapter we describe a study of routinely collected, computerised NHS data exploring these potential benefits among patients referred to the IAPT demonstration sites. The collection and analysis of data were carried out by

the Clinical Informatics Group, at St George's University of London, led by Simon de Lusignan, supported by the University of Sheffield team, who introduced them to the practices and continued to liaise between the two.

Routinely collected GP data had previously been collected and analysed by the St George's team, to try to measure the impact of IAPT at the level of the practice population. Computerised prescribing data from 25 GP surgeries (23% of the 110 which could potentially refer their patients to the IAPT demonstration services) were collected using Morbidity Information Query Export Syntax (MIQUEST) which is Department of Health sponsored data interrogation software, used to extract and pool information from a variety of GP computer systems. Data were extracted for the first two quarters of the four calendar years 2004 to 2007 (before the introduction of the IAPT services) and compared with prospective data gathered for the first two quarters of 2008 (post their introduction).

These data showed a year-on-year rise in antidepressant use between 2004 and 2007 which continued after the introduction of the IAPT services [85]. However, this finding has to be viewed against a background of steadily rising antidepressant prescribing nationally [86]. It was therefore possible that any benefit of reduced prescribing among patients referred to the IAPT services was outweighed by a rise in prescribing among patients not referred.

It had not been possible to identify patients who had been referred to the IAPT services in the previous study, in order to determine differences in medication use in those patients specifically. However, in the study reported here, new techniques were developed by the St George's team to enable linkage of data on patients referred to the IAPT services with their general practice prescribing data. Furthermore, it proved possible to examine those patients' use of secondary care services for physical symptoms, through linking their IAPT and GP practice data to data from secondary care computer systems.

7.2 Research questions

The routinely collected data from GP computer systems, IAPT, and secondary care were collected as an add-on research study to the original research study. This add-on study was ethically approved (UCL/UCLH Committees on the Ethics of Human Research - Committee Alpha. Ref: 08/H0715/101 15/10/08) and as there was transient holding of personally identified data during the encryption process also by PIAG (Patient Information Advisory Group Ref: PIAG 6-06(h)2008 17/12/08) for Section 251 support. The research was sponsored by Newham PCT, and approved

by local research governance committees of: Newham and Doncaster PCTs and the East London NHS Foundation Trust.

Firstly, the study included an assessment of any age, gender, social deprivation or ethnic minority bias in the referral of patients with common mental disorders to the IAPT services.

Secondly (and this was the study's primary objective), the study determined whether differential consumption of health service resources was associated with referral to the IAPT demonstration services for people with anxiety and depression.

The health service resources included the use of psychotropic medications, and the use of secondary care services for physical health problems, including outpatient, in-patient and accident and emergency services. Levels of utilisation among the patients referred to the IAPT programme were compared with levels for control patients with anxiety and depression that were not referred, matched for age, gender, and practice.

7.3 Summary of method

7.3.1 General practice data

All practices in the two demonstration sites were approached and asked to provide written consent to participate. To be eligible for the study, each practice had to have an appropriate computer system from which data could be extracted using MIQUEST, to have had the same computer system for the previous five years, and no plans to change it during the study period.

MIQUEST when written in its 'Remote' mode creates a 'one line per patient' comma separated flat file with each line uniquely identified by a pseudo-anonymised patient identifier - the MIQUEST ID. The GP practices retain the MIQUEST generated ID, which allows the practices, and only the practices concerned, to identify the individual patients as appropriate. In its 'Local' mode it allows the extraction of patient identifiable data such as the full postcode, date of birth etc. All data are anonymised before being extracted from practices for external analysis, but retain a pseudonym for each patient so they can be analysed at the individual patient level (i.e. they are 'pseudonymised').

The following routinely collected patient data were collected from the practice computer systems:

• Person identifiers for the linkage and subsequent de-identification process, specifically name, date of birth, and postcode

• Demographic information: age, gender, ethnicity, registered date

- Diagnoses of common mental health problems including all codes for depression or anxiety:

The READ codes for affective (depressive) disorders searched were E2B; E2B1; 1B1U; E112%; E204; 1465; E2003; E113; E11z2; E135; E291; E2B1; Eu32%; Eu33%; Eu34%; Eu412; 1BT; and 1B17.

Codes searched for other neurotic, stress and somatoform disorders (including generalised anxiety disorder, panic disorder, 'stress', and 'tension' etc.) were E28%; 1B1L; 1B1T; 13JM; 13HT1; 67J%; 9ON%; R00zW; Eu4; Eu40; Eu40z; E2001;1B1V; Eu400; Eu41%; E2021; E2002; R2y2; 1B12; Eu43%; Eu46%; E2781; F2626; Eu454; R040%; 1B1G; 1BB%; and 1BA2-1BA8.

- Sickness certificates issued by GPs and recorded in the medical record.

- Prescriptions of psychotropic medications

- Referral to further care

- Deprivation: Postcode was transiently extracted and linked to deprivation index. : Full postcodes were extracted separately in 'Local' mode, but immediately transformed into deprivation scores within GP computer systems using the Multiple Deprivation Index. The interim local file generated by MIQUEST with full postcodes was then deleted from the GP computer systems. Only the Deprivation scores along with the first half of the postcode to establish general location were retained for analysis.

- Long-term physical health problems: To estimate co-morbid physical problems among the patients with common mental health problems, diagnostic data were extracted for: diabetes mellitus; chronic kidney disease (CKD); heart failure; ischaemic heart disease (IHD), hypertension; and chronic obstructive pulmonary disease (COPD).

7.3.2 IAPT service use data

The IAPT Programme adopted a Minimum Data Set (MDS) for its outcome framework. This allowed routinely collected data for all adults referred to the IAPT Programme in the two demonstration sites between October 2007 and September 2008 inclusive to be included in the analysis. This period was chosen as it allowed more than 12 months after the start of the services for referral rates to reflect ongoing activity among incident cases rather than the initial higher rate of activity expected due to the build-up of patients within practices with more long-standing problems (prevalent cases), for whom no service had been previously available.

Service activity data in IAPT were routinely collected including the date of referral, date of assessments, the nature of the interventions offered, the number of sessions in each step of the stepped care pathway, the numbers who dropped out from therapy, and the global outcome measures of anxiety and depression. Patients' names, postcodes, and dates of birth were used for data linkage between the IAPT service data and the practice data, then the data were subsequently de-identified within the service premises. Again, all data were pseudonymised and encrypted before being removed for external analysis.

The IAPT programme did not include the NHS unique identifier (NHS number) as it was not made part of the IAPT MDS. The lack of NHS number meant that matching of IAPT to SUS (secondary care use of services) and GP data had to be carried out using other identifiers. We had separately to pseudonymise strong identifiers, encrypt them and then match using fuzzy logic: a new technique called SAPREL – Secure and Private Record Linkage. This linkage had to cope with potential differences of use of forename, e.g. Liz v. Elizabeth; second name – either due to use of married or other surname or sometimes a shortened form of surname, and any transcription errors which might occur. The SAPREL process is new and unique – using fuzzy logic to link cases without the researcher knowing the identity of the patient. We intend to report experience with this process separately.

7.3.3 Secondary care use of services (SUS) data

NHS Hospital Trusts routinely collect data about patient care in outpatient clinics, day treatment units, and as hospital in-patients. These data are organised differently from GP data. In primary care each individual encounter is usually recorded using a combination of coding and narrative text. However, in secondary care all encounters relating to a particular problem are grouped together using an overarching term to bring together all the individual events that contributed towards a particular "spell" of care. Access to SUS data was negotiated through the lead practitioners and managers of the primary care trusts (PCTs) of the two demonstration sites, and the PCT data guardians provided written consent.

Raw Secondary Uses Service (SUS) data for the period between 01/10/2007 and 30/04/2009 were accessed through the Information Services of the Primary Care Trusts (PCTs) in the 2 study sites. A subset of the cleaned data was extracted to contain only records with NHS numbers which could be matched to an existing NHS number in the participating practices' data. Service activity data used for the study included routinely collected data on out-patient clinics (OP), Accident and Emergency (A&E) attendance, day treatment units, and hospitals, for in-patient (IP) care episodes. Data fields

collected for the study were dates, diagnoses and procedures relating to the care episodes in these secondary health care settings, to infer utilisation of secondary healthcare OP, A&E, and IP services. Once again, patient identifiable information for data linkage, specifically name, postcode and date of birth, were kept within the premises where such data were held or accessed, and all data were pseudonymised and encrypted before being extracted for external analysis.

SUS data do not include any data from mental health services. The IAPT clinics have their own proprietary systems; and these are entirely separate from other mental health records. Mental health services are the last part of the NHS to be computerised. The predominant system used in London is CSE Servelec's RiO system. This has structured windows with free-text (e.g. one for risk assessment, one for formulation, etc.) and does not have coded data in the way that hospital SUS data and GP system data do. It is not yet practical to search free-text mental health records (and possibly unlikely that researchers would get permission to do so without patients' individual consent).

7.3.4 Data linkage

Specific approval from the Patient Information Advisory Group (PIAG) under Section 251 exemption was obtained to support the use of patient identifiable information for the purpose of linkage using SAPREL, followed by de-identification and pseudonymisation.

Data files from the 3 sources were linked using SAPREL, a process which PIAG (Patient Information Advisory Group) commended as an example of best practice.

The final linked dataset was imported into the Statistical Package for Social Sciences (SPSS) for analysis, after vectorisation. Vectorisation is the process of turning scalar data into vectors (to enable them to be processed by applications like SPSS). The non-GP data in this project (from SUS and IAPT clinics) were output as datasets linked to a unique ID (in our case the NHS number). They were vectorised into a single row for each patient in the final data table. SUS data were output as a minimum dataset for each type episode (e.g. for in-patient length of stay, specialty, procedure etc.). This same dataset was output for each stay an individual patient had in hospital, or attendance at A&E or out-patients. The same applied to each attendance for patients who attended the IAPT clinic more than once. The vectorisation process took the first dataset for each category (IP, OP, A&E and IAPT attendances) and converted this dataset into a row so that the dates for a first attendance of the same type were all in the same set of columns; those for the second likewise and so on. This vectorisation process made the data table very large as the same minimum dataset was issued for each

attendance and some of the study patients had attended as A&E, OP and IP cases more than 20 times.

7.3.5 Analysis

To determine any association between health service resource consumption and referral to IAPT services, the patients identified as having been referred to IAPT were compared with control patients who were identified as having received a diagnosis of a common MH disorder after 1st April 2007, matched for age, gender, and practice. The ratio of cases to controls selected was 1:6.

7.4 Findings

7.4.1 Data overview

7.4.1.1 GP data

A total of 20 GP practices in the 2 study sites participated in the study, with a combined population of 152,302 patients:

• Doncaster: 10 practices with a total list size of 73,670, of whom 59,349 were aged 16 or over

• Newham: 10 practices with a total list size of 78,632, of whom 61,850 were aged 16 or over

This represents 18% of the total number of 110 practices eligible to refer patients to IAPT (46 in Doncaster and 64 in Newham).

7.4.1.2 IAPT data

We collected 5581 records from the two IAPT project teams for the period between 1st October 2007 and 30th September 2008:

• Doncaster: 4,026 records

• Newham: 1,555

Of these, 1,153 IAPT records (with 1,118 unique patients) were linked to GP data from the participating practices. These services treated a further 4,429 people, but of course we could not link their records to data from GP practices as their GPs had declined to participate in this part of the study.

7.4.1.3 Secondary Uses Service (SUS) data

The Doncaster & Newham counts prior to vectorisation were: -

- A&E records: Doncaster 22,277; Newham 25,498
- In-patient records: Doncaster 25,316; Newham 25,653
- Out-patient records: Doncaster 119,358; Newham 129,337

The final data file had 152,328 patient records and over 9,800 variables, with over a billion cells in the data table. There were 26 duplicated pseudonymised NHS numbers, generated when patients changed their GP practice or their address (as indicated by the same gender and year of birth, but different Practice ID codes or different pseudonymised postcodes). Having combined data for each of the 26 duplicates, our total file for analysis included 152,302 patients.

7.4.2 Recorded diagnoses of mental health problems

Within the GP computer data diagnoses of affective disorders included:

- Manic episode

- Bipolar affective disorder

- Depressive episode

- Recurrent depressive disorder

- Persistent affective disorder

- Other and unspecified affective disorder

- Mixed anxiety and depressive disorder

- History of depression

Diagnoses of neurotic, stress and somatoform disorders included:

- Panic disorder

- Generalised anxiety disorder

- Mixed anxiety and depression

- Reaction to severe stress & adjustment disorders

- Dissociative disorders

- Unexplained somatoform complaints

- Other neurotic disorders

Obsessive compulsive disorders (OCD), which has a separate code, were also extracted using MIQUEST.

7.4.2.1 Common mental health disorders

A new variable was derived which combined all recorded diagnoses of affective disorders, neurotic disorders or OCD, into one diagnosis of 'common mental health (MH) disorder'. As some of the diagnoses of MH problems in GP information systems went back many years, a limit to the time period for these diagnoses was placed on the data for the main analyses. 'Recent diagnoses' of common mental health disorders were limited to recorded diagnoses in GP information systems on or after 1st April 2007 – the start of the data collection from the IAPT service.

Table 42 below shows the numbers of patients identified with affective, neurotic and obsessive-compulsive disorders in the participating practices, as a percentage of the total combined number of 152,302 identified patients.

Table 42 below also shows that 17.8% of the population, or 22.3% of the adult population (aged 16 or over) had received a diagnosis of a common MH disorder since the practices had started recording records on computer, but since 1st April 2007 the total 12-month prevalence was just over 8%, or just over 10% in the adult population.

The proportions among individual classes of disorder, of 7.2% for affective, 4.1% for non-affective neurotic disorders, and less than 0.1% for OCD, add up to more than the total proportion of 10.1% because individual patients could have received more than one class of diagnosis during the period for which data were extracted.

Table 42. Frequencies of affective disorders, neurotic, stress and somatoform disorders and common MH disorders in the participating GP practices

Class of common MH disorder	Number with ever recorded diagnosis	Number with diagnosis recorded after 1st April 2007
Affective disorders	20,982 (17.3%)*	8,721 (7.2%)*
Neurotic, stress and somatoform disorders	12,046 (9.9%)*	4,984 (4.1%)*
OCD	144 (0.1%)*	59 (<0.1%)*
Total Common MH disorders	27,017 (22.3%)*	12,210 (10.1%)*

* Percentage of the total combined number of 121,199 identified patients aged 16 and over.

The practices in Doncaster recorded a higher proportion of people with common MH problems than the practices in Newham. The total prevalence of diagnoses since 1st April 2007 ranged from 11.2% to 16.8% among the 10 Doncaster practices, compared to a range from 1.9% to 9.1% found among the 10 Newham practices.

7.4.2.2 Long-term physical health problems

The recorded prevalence of long-term physical health problems in the total study sample were: diabetes mellitus 5.7%; chronic kidney disease (CKD) 4.6%; heart failure 0.8%; ischaemic heart disease (IHD) 3.8%, hypertension 14.5%; and chronic obstructive pulmonary disease (COPD) 11.8%.

Patients with common MH problems were a little more likely to have co-morbid long term conditions than the total sample, but apart from COPD the differences were not large (diabetes 6.3%; CKD 5.3%; heart failure 0.9%; IHD 4.7%; hypertension 15.6%; and COPD 18.5%).

7.4.3 Referral of patients with common MH disorders to IAPT

Table 43 below shows the numbers and percentage of patients, with the various types of recorded diagnoses of common MH disorders within the 20 participating practices, who were referred to the IAPT services during the period of data extraction 1st October 2007 to 30th September 2008.

The table shows that overall only 6.3% of patients with a recorded diagnosis of a common mental health disorder were referred to the IAPT services by the participating practices during the year in question. Furthermore, the proportion of patients referred varied by diagnosis, being slightly greater for affective disorders (7.5%) than non-affective neurotic disorders (5.6%), with the proportion of patients with OCD who were referred being possibly greater still (8.5%), although it is difficult to draw conclusions when so few patients in that group were referred.

Table 43. Referral of adult patients (16 years and over) with common MH disorders to the IAPT service

Recent recorded diagnosis (after 1st April 2007)	Number not referred to IAPT	Number referred to IAPT	Total number (100%)
Affective disorders	8,070	651 (7.5%)	8,721
Neurotic, stress and somatoform disorders	4706	278 (5.6%)	4,984
OCD	54	5 (8.5%)	59
Total Common MH disorders	11,436	774 (6.3%)	12,210
No recorded diagnosis of a common MH disorder	108,645	344 (0.3%)	108,989

The table also shows that 344 patients, 30.7% of the total of 1,118 who were referred to the IAPT services, had no recorded diagnosis of a common mental health disorder within the practice after April 1st 2007. The GP records of these 344 individuals were further examined for the presence of other mental health problems such as substance misuse, psychosis, sleep

problems etc. and for the prescription of psychotropic drugs including antidepressants, anxiolytics and hypnotics. The large majority of these 344 did not have a recorded substance misuse, psychosis, or sleeping problem; only 12 had a recorded diagnosis of alcohol misuse, two had a diagnosis of misuse of opioids/cannabinoids; four had received a diagnosis of psychosis; and 30 had been seen with sleeping difficulties. A total of 146 had been prescribed antidepressants, 31 anxiolytics, and 39 hypnotics.

This suggests a proportion of patients were seen with mental health problems and may have been referred to IAPT without the GP or other health professional entering a specific READ coded diagnosis. The alternative explanation is that the patients referred themselves to the IAPT service, or were referred by an alternative service, and were not seen and diagnosed within the practice at all. Self-referral was encouraged in Newham where stigma was associated with seeking help from the GP.

It should also be noted that, of the 1,118 patients referred to the IAPT services, only 723 (64.7%) received an initial brief assessment, only 530 (47.4%) had a full initial assessment, and 588 (52.6%) went on to receive treatment by the services.

7.4.4 Referral rates by age, gender, and ethnicity

To explore possible bias in the referral of patients to IAPT services, variation in the proportion of patients with common MH disorders who were referred was examined by recorded age, gender, and ethnicity (Table 44).

Table 44. Referral of adult patients (16 years and over) with common MH disorders to the IAPT service

Demographic factor	Patients with common MH disorders not referred to IAPT N = 11,436	Patients with common MH disorders referred to IAPT N = 1,118	Statistical significance of differences (chi-square test)
Older age (% over 65 years)	1290 (11.3%)	17 (1.5%)	p < 0.001
Gender (% female)	7687 (67.2%)	542 (48.5%)	NS

Ethnicity (% black and ethnic minority patients)	1,441 (12.6%)	97 (8.6%)	p < 0.001

7.4.4.1 Age
The rate of referral to IAPT services was much lower among people with common MH disorders aged 65 and over. This pattern was similar for both Doncaster and Newham.

7.4.4.2 Gender
More males than females were referred to the IAPT services. Again, this pattern was similar for both Doncaster and Newham. However, among only those patients who had received a recent diagnosis of common MH health problems, the gender difference in the rate of referral was much smaller. The pattern was also different between Doncaster and Newham: in Newham, more males than females were referred to the IAPT services among those with a recent diagnosis of common MH health disorders.

7.4.4.3 Ethnicity

People from black and minority ethnic groups were less likely than whites to be referred to IAPT services. Again, this pattern was similar for both Doncaster and Newham. However, among those with a recent diagnosis of common MH health disorders, the difference in the rate of referral by ethnicity was much smaller.

7.4.4.4 Long-term physical health problems
There were no significant differences in the proportion of people with long-term physical health problems between the IAPT referred and control groups.

7.4.5 Sickness certification

The number of sickness certificates recorded in the GP practice computer systems was determined for all patients with a diagnosis of a common MH disorder, and the mean number for those referred to the IAPT services (n = 744) was compared with the mean for six control patients for each case, matched for age, gender, and site, who had received diagnoses of common MH disorders after 1st April 2007 but had not been referred to IAPT services (n = 4,464).

The mean number of certificates recorded for the IAPT referred patients was 7.29, compared to 5.93 for the controls. The mean difference was therefore 1.36 (95% confidence interval 0.89 to 1.82), which was statistically significant (t-test for equality of means, 2-tailed significance $p < 0.001$).

It should be noted that the period of recording includes a period of time before referral to the IAPT services for the cases, as well as the time period recorded after referral. It is not possible therefore to tell from this analysis whether on the one hand referral was associated with a subsequent increase in sickness certification, or on the other referral was more likely among patients already receiving greater numbers of certificates.

In order to determine the timing and direction of association it was necessary to examine whether treatment in the IAPT service was followed by an increase in the mean number of sickness certificates. Therefore a further analysis was carried out, looking at changes in certification before and after treatment of the IAPT referred patients (see section 7.4.8 below).

7.4.6 Use of psychotropic drugs

The data extraction software MIQUEST is limited in that it only allows the extraction of up to 15 prescriptions of any one type of medication going forward from the start of the patient's computerised record, and up to 15 prescriptions looking backwards from the end of the record. This means that it cannot gather all prescriptions for patients with more than 30 in total. In our dataset 4530 patients (3%) were found to have 30 or more records of anti-depressants, 423 patients (0.3%) 30 or more records of anxiolytics, and 648 patients (0.4%) 30 or more records of hypnotics. The data below therefore may underestimate the total number of prescriptions provided to around 3% of patients included.

Table 45 below shows the numbers of prescriptions of antidepressants, anxiolytics and hypnotics recorded in the GP practice computer systems for patients with common MH disorders, comparing the 744 patients referred to the IAPT services with the 1,259 controls.

Table 45 shows that being referred to IAPT was associated with the recording of significantly greater numbers of prescriptions for these three classes of psychotropic drugs. There was a trend in the same direction across each of the three classes but this was statistically significant only for the antidepressants. The numbers receiving anxiolytics and hypnotics were smaller, and so the lack of statistical significance of differences in those groups may have been due to the smaller sample size (a 'type II error').

Table 45. Prescriptions of psychotropic drugs for patients with common MH disorders referred and not referred to the IAPT services

Type of psychotropic medication prescribed	Mean no. of prescriptions among cases referred to IAPT (n = 744)	Mean no. of prescriptions among controls not referred (n = 1,259)	Mean difference (and 95% confidence interval)	Significance of the difference (2-tailed t-test)
Anti-depressants	7.02	5.71	1.30 (0.77 to 1.84)	p < 0.001
Anxiolytics	0.55	0.48	0.07 (-0.11 to 0.25)	p = 0.440
Hypnotics	0.56	0.60	-0.04 (-0.24 to 0.16)	p = 0.693
Total for all three classes of drugs	8.14	6.80	1.33 (0.67 to 2.00)	p < 0.001

Once again, the period of recording includes a period of time before referral to the IAPT services for the cases, as well as the time period recorded after referral. It is not possible therefore to tell from this analysis whether on the one hand referral was associated with a subsequent increase in psychotropic drug prescribing, or on the other referral was more likely among patients already receiving greater numbers of prescriptions. Once again, a further analysis was carried out, looking at changes in antidepressant prescriptions before and after treatment of the IAPT referred patients (see section 7.4.8 below).

7.4.7 Use of secondary care health services

Table 46 below shows the numbers of inpatient, outpatient, and Accident and Emergency episodes, plus the total number of days spent in inpatient stays, from SUS data for patients with common MH disorders. Once again the 744 patients referred to the IAPT services are compared with the 1,259 controls.

Table 46 shows that being referred to IAPT was associated with receiving significantly fewer inpatient, outpatient, and A&E episodes of secondary care, and significantly fewer days spent as an inpatient, for physical problems. In particular, it is noteworthy that the number of days spent as an inpatient was approximately half among the IAPT referred patients.

Table 46. Use of hospital services for physical health problems among patients with common MH disorders referred and not referred to the IAPT services

Hospital resource use	Cases referred to IAPT (n = 744)	Controls not referred to IAPT (n = 1,259)	Mean difference (and 95% confidence interval)	Significance of the difference (2-tailed t-test)
No. of A&E attendances	0.80	1.38	-0.59 (-0.78 to -0.39)	$p < 0.001$
No. of outpatient episodes	3.72	6.11	-2.39 (-3.04 to -1.73)	$p < 0.001$
No. of inpatient episodes	0.68	1.68	-0.99 (-1.18 to -0.80)	$p < 0.001$
Total no. of days spent in IP care	1.56	4.23	-2.68 (-4.11 to -1.24)	$p < 0.01$

Once again however, it should be noted that the data shown relate to the total period of recording of SUS data which includes a period of time before referral to the IAPT services for the cases, as well as a time period recorded after referral. It is not possible therefore to tell from this analysis whether on the one hand referral was associated with a subsequent reduction in the use of secondary care services, or on the other referral was more likely among patients already receiving less in the way of secondary care for physical problems. Once again, a further analysis was carried out, looking at changes in use of hospital services before and after treatment of the IAPT referred patients (see next section 7.5.8).

7.4.8 Before-and after analysis of sickness certification, prescribing, and use of secondary care services

7.4.8.1 Model used for analysis

In order to determine the timing and direction of the associations described above it was necessary to examine whether any treatment within the IAPT services of patients referred to them was followed by a change in prescribing, sickness certification, or use of hospital services, when compared to control patients with common mental health disorders who were not referred to IAPT.

In order to do this, an estimate had to be made about the timing of treatment of the patients referred to IAPT, as it was too difficult to relate changes in prescribing, sickness certification and secondary care service use to the exact start and end dates of treatment within IAPT for individual patients, given the complexity of the available computerised data. We therefore conducted the before and after analysis using data for six months either side of the estimated date of initial assessment of the patient by the IAPT service. The assessment date was estimated by taking the date of referral and adding 44 days, the mean delay from referral to assessment.

An intention to treat approach was adopted, including all patients referred even if they did not actually attend for assessment and treatment. We excluded patients with an estimated assessment date less than six months before the date of data extraction (30th September 2008), so that a full six months of data after assessment could be included, for comparison with the six months before. This had the effect of reducing the total number for analysis, shown in the tables below. The 3% of patients with more than 30 prescriptions were also excluded from the analysis of antidepressant prescribing. This had the effect of further reducing the total number.

7.4.8.2 Findings

Table 47 below shows that total antidepressant prescriptions increased by more among the IAPT referred patients than among the controls, but it should be noted that the number of people prescribed antidepressants fell from 500 (48%) to 442 (43%) in the IAPT group, while the number in the control group rose from 428 (9%) to 547 (12%); chi square p< 0.001.

Table 47 also shows that there were comparative reductions in sickness certificates, hospital admissions, occupied hospital bed-days, out-patient appointments and A&E visits among the IAPT referred group. Table 48 shows that these differences were statistically significant for the reduction in sickness certificates, and just significant for the reduction in Accident & Emergency admissions.

Exploratory analysis restricted to the people who actually attended for treatment showed trends in the same direction as above, but not statistically significant differences.

Exploratory comparison of changes in service use between patients who had had common mental health problems for at least 6 months and those with a shorter duration of symptoms showed no differences between the groups.

Table 47. Changes in antidepressant prescribing, sickness certification, and use of health services among adult patients (aged 16 and over) with common MH disorders referred to the IAPT services compared with controls (adjusted for age, gender, and practice)

Resource		n	Pre referral mean	Post referral mean	Absolute change from baseline	Difference in absolute change from baseline:
Antidepressant	IAPT	1118	1.27	1.65	0.38	0.15
prescriptions	Control	5286	0.54	0.77	0.23	
Sickness	IAPT	1118	0.49	0.39	-0.1	-0.11
certificates	Control	5286	0.18	0.19	0.01	
Hospital	IAPT	712	0.21	0.19	-0.02	-0.04
admissions	Control	3072	0.38	0.41	0.03	
Total days	IAPT	712	0.53	0.26	-0.27	-0.25
admitted	Control	3072	0.80	0.78	-0.02	
Outpatient	IAPT	712	0.96	1.09	0.13	-0.16
attendances	Controls	3072	1.37	1.66	0.29	
A&E	IAPT	712	0.25	0.23	-0.02	-0.12
attendances	Controls	3072	0.39	0.49	0.1	

Table 48. Significance of changes in antidepressant prescribing, sickness certification, and use of health services among adult patients (aged 16 and over) with common MH disorders referred to the IAPT services compared to controls (adjusted for age, gender, and practice)

Resource		n	Mean change	Differ ence	SE differ ence	95% CI of the difference	t-test Sig. (2 tailed)
Antidepressant prescriptions	IAPT	1118	-0.38	-0.15	0.07	-0.29 to -0.02	*p*=0.028
	Control	5286	-0.23				
Sickness certificates	IAPT	1118	0.1	0.11	0.03	0.04 to 0.17	*p*=0.002
	Control	5286	-0.01				
Hospital admissions	IAPT	712	0.02	0.04	0.04	-0.03 to 0.11	*p*=0.229
	Control	3072	-0.03				
Total days admitted	IAPT	712	0.27	0.25	0.29	-0.31 to 0.82	*p*=0.385
	Control	3072	0.02				
Outpatient attendances	IAPT	712	-0.13	0.16	0.12	-0.02 to 0.35	*p*=0.081
	Control	3072	-0.29				
A&E attendances	IAPT	712	0.02	0.12	0.04	0.06 to 0.19	*P*<0.001
	Control	3072	-0.1				

7.5 Key findings and their implications

We found an overall annual prevalence of diagnoses of common MH disorders of 10% among patients aged 16 and over. This may be compared with findings from the 2007 Office of National Statistics household survey of adult psychiatric morbidity in England [87] which found a one-week prevalence of 16.2% among 16-64 year olds. It has been shown that only around half the cases of common MH problems presenting to GPs are recognised [88-90], so a 10% prevalence of recorded diagnoses among adult patients is consistent with previous research.

It is striking that the range of levels of diagnoses in Newham practices were generally significantly lower than the range in Doncaster. It is unlikely that common MH disorders are actually less common in Newham, given it is one of the most deprived areas in the country, and this is likely to represent considerable under-diagnosis or under-reporting of problems. Under-reporting may in turn be related to the acknowledged local difficulties in delivering mental health services in Newham.

The fact that a relatively small proportion of people with common MH disorders (6.3%) were referred to the IAPT services for intervention means that any effect of IAPT at the population level is likely to be diluted.

However, this still compares quite favourably to the proportion of patients likely to receive referral for psychological treatment in areas where IAPT services do not exist, of around 1% [83,91].

It should also be noted that, of the 1,118 patients referred to the IAPT service, only 723 (64.7%) received an initial brief assessment, only 530 (47.4%) had a full initial assessment, and only 588 (52.6%) actually got as far as receiving treatment. This was partly due to patients being diverted to other, more appropriate services, as signposting was a significant function, particularly of the Newham service, which signposted 20% of patients to a sister NHS service.

Generally, there was little evidence of referral bias according to gender or ethnicity. There was a bias against referring patients aged 65 and over, which may be because one of the selling points of the IAPT services is that they can help patients return to work, which might well lead to a bias towards referring people of working age, even where no specific age criterion for acceptance of referrals exists, as in the pilot sites. The Newham IAPT team did consider that there was a bias against referring patients of ethnic minorities, and that reducing this bias was a major task undertaken by the team.

The patients referred to IAPT were more likely to be receiving sickness certificates and antidepressants than control patients with similar diagnoses, which suggests there was a greater severity of mental health problems in the patients referred to IAPT. This would be expected, as antidepressant treatment and referral are more likely among patients with a greater severity of depression [91].

The mean number of prescriptions of antidepressants increased in both IAPT and control groups after treatment of the IAPT group, so referral to IAPT did not bring down overall antidepressant prescribing, although the proportion of people taking antidepressants in the IAPT referred group did reduce by 5% while the proportion increased in the control group by 3%. Even this small reduction in proportion may be seen as encouraging, as there is no systematic method through which patients referred to IAPT services receive advice or guidance on stopping or reducing their drug treatments. Psychiatric advice about medication is not part of the basic IAPT model, and few services have so far established explicit links with psychiatrists [81]. In Newham all IAPT referred patients could be referred by the therapists for expert psychiatric advice on an as required basis, although in many cases the action taken was to optimise medication rather than discontinue it, continuing it beyond the recovery period in order to maintain improvement in line with good practice guidelines.

Our results suggest that some IAPT patients discontinued their antidepressants as a result of referral to the IAPT service, which may have been due to a psychiatric review associated with the referral, at least in Newham. Other patients were likely to have been judged to need antidepressants and encouraged to continue them for longer. Other research suggests there is a general trend to prescribe antidepressants in general practice for longer periods, which was reflected in the increase in the total number of prescriptions of antidepressants per patient observed over time in both the IAPT and control groups.

Sickness certification reduced among the IAPT referred patients after referral but increased slightly among the controls during the same time period, and this difference was statistically significant. This is in line with other findings in relation to the IAPT services [36].

Attendances at Accident & Emergency also reduced among the IAPT referred patients while staying the same for controls, and again this difference was just significant statistically. There was also a reduction in inpatient days among the IAPT referred group which didn't quite reach statistical significance, which may have been due to a lack of power to find a significant difference, as the numbers of patients who were admitted was relatively small (only 155 in the IAPT referred group). However outpatient attendances increased among the IAPT referrals slightly more than they did among controls.

This study has important limitations. Routine data have shortcomings: we can only extract "coded" data – not the text record. However, we are aware of routine data strengths and shortcomings [92]. Our "after" period of analysis may not have been long enough for changes in hospital utilisation to work through. Length of stay and A&E attendances may be more amenable to change than out-patients and inpatients bookings which may have been made up to 18 weeks before. The study may have been underpowered.

7.5.1 Conclusions

Overall, our findings suggest that IAPT referral is being appropriately targeted on people with a greater severity of problems, and is reducing sickness certification, in line with the IAPT services' own data.

However, we found no evidence to suggest that referral has a significant effect on overall antidepressant prescribing, although the proportion of

patients prescribed antidepressants did decrease by 5% in the IAPT referred group.

There is some indication that referral to IAPT might lead to a reduction in the use of secondary care non-mental health services, particularly in Accident and Emergency attendances. It is unfortunate that computerised data on the use of secondary care mental health services are not available in a form which would allow similar analyses to those we carried out for physical health care services, and more research is urgently needed on the effects of IAPT on referral to mental health services.

These changes represent significant potential benefits in economic terms, at least among the patients referred. It would be difficult to estimate the overall economic benefit given the lack of individual patient data on use of services. However, the fact that only 6% of patients with common mental health problems were referred means that any overall effect is likely to be significantly diluted.

Project 08/1610/154

8 Discussion and conclusions

This final chapter summarises the key findings of our evaluation of the IAPT demonstration sites and discusses these in relation to the specific aims of the IAPT programme. We then outline the limitations of this evaluation, discussing the tensions intrinsic to commissioning and undertaking evaluative research of this nature in a fast-moving area of policy and make some recommendations for how this could be improved. Finally, we provide some advice on what can be learned from the demonstration sites for the future development of IAPT.

Overview and commentary on key findings

8.1.1 Service delivery

Although the services in Doncaster and Newham both aimed to improve access to psychological therapies, they were grounded in very different local organisational and socio-demographic contexts and their aspirations and service models differed accordingly. For this reason our evaluation design did not seek to make direct evaluative comparisons between the two sites, but rather to evaluate each within its own context.

Doncaster succeeded in providing a high-volume service for people with depression and anxiety disorders across a wide range of severity which gave local people unprecedented access to guided self help informed by cognitive behavioural principles and, to a lesser extent, cognitive behaviour therapy.   In opting for this delivery model, over the 35 month audit period, the Doncaster service was able to accept an average of 322 referrals per month, an average of nine per month accessed CBT and an average of 84 interventions were completed per month. The original vision for partnership working between NHS and non-NHS agencies including cross-referral from employers, occupational health services and job centres was not realised.

Newham succeeded in establishing a cognitive behaviour therapy service accessible to primary care patients and to self-referrals. It engaged people of diverse ethnicity in a socially disadvantaged area, within a complex matrix of existing services, working alongside employment coaches and systemic therapists. It then succeeded in adapting the original service model to accommodate a greater volume of referrals through offering more

guided self-help including computerised CBT, alongside access to CBT. This 'mixed' model approximately balanced the numbers receiving low and high intensity interventions. Taking an average across the 33 month audit period, the service accepted an average of 101 referrals per month, 20 per month accessed CBT, and 18 treatments per month were completed. Of those referred, 43% went on to have an intervention; those who did not were often 'signposted' to other services, and a proportion was re-referred later. Newham went to considerable lengths to engage a 'hard to reach' population through a system of assertive follow-up and flexible engagement. Over half of referrals were people of non-White ethnicity and rates increased over the audit period. The aspiration to make the service accessible to self-referral through community liaison was realised, with self-referrals making up nearly a quarter of referrals.

Both services kept waiting times for first contact short and in both services under half of people referred received an intervention, defined as two or more contacts with the service, tending to be those with more severe depression and anxiety. Those not engaging with the service were provided with advice or 'signposted' to other services and a proportion was re-referred later. Most people had suffered significant difficulties for more than a year; a quarter to a third of people presented with severe levels of depression and just under a half had severe levels of anxiety. The service was used by unemployed people and people on benefits, who tended to have more severe difficulties.

8.1.2 Organisational issues in implementation

We were aware of the enormous challenge to the organisations implementing this complex new model of service within a very short timescale. Whilst our evaluation had a three year timescale, the services were being asked by the Department of Health to demonstrate effectiveness within six months of accepting their first referrals, in order to prepare the business case to Treasury for roll-out of the programme in the following year. Our organisational case study addressed the key issues for organisation and delivery of services through a qualitative analysis of interviews with key informants within and outside the services, at two points in the implementation process. Seven broad themes emerged to provide the framework for our analysis, andtwo overarching themes, that of partnership working and local sustainability. We shall not attempt to summarise the themes here, but rather shall select some salient points. The first of these emphasised the unusual nature of the external pressures that people perceived to be operating in relation to the IAPT programme, which had a high profile within the Department of Health and more generally. The shift from national demonstration project to locally

commissioned service provided relief from these pressures, and a sense of greater freedom in implementation.

The aspirations of IAPT were clearly in terms of addressing mental health inequalities, breaking down barriers and creating a service that supported individuals in remaining active in society and in their community. Moving away from traditional clinical delivery methods was cited as a key way of avoiding some of the stigma attached to mental health issues, overcoming shortfalls in support for those in employment who are finding it difficult to cope and in providing access to those in previously hard to reach communities i.e. those from black and ethnic minority communities and non English speakers. Discussion of the innovation itself with key stakeholders suggests the centrality of relationships with partners in getting the new services up and running and awareness with hindsight of the need for more time to enable these partnerships to develop. The complexities of inter-organisational partnerships within IAPT were highlighted, including who is 'in' or 'outside' the partnership, who 'owns' it, dynamics between primary and secondary care, the quality of relationships and the extent of collaboration. Achieving a genuinely seamless pathway by good collaboration between primary and secondary services was an enduring difficulty at both stages of data collection and requires continuing attention. There was more dissatisfaction than satisfaction expressed with the way the partnerships were working and recognition that there is a considerable difference between partners signing up to working collaboratively at the strategic level (e.g. developing the shared vision) and making it work operationally on a day to day basis. Interviewees identified a number of factors or activities that were found helpful in overcoming distrust, misunderstandings and practical barriers between partners and facilitating genuine partnership.

The time required for genuine partnership working is considerable, and within an NHS context, it could benefit from external facilitation (e.g. commissioners), commitment at the top of the organisations, plus individuals skilled in organisational development in both the organisations needing to collaborate – these are the opinion leaders, change agents and 'boundary spanning' individuals [10].

Some of the factors that could hinder partnership can be considered. For example, asking NHS Trusts who need to work in partnership to enter into competitive bidding for commissioned services seemed to undermine collaboration across the care pathway (although this did not seem to pose a problem in other settings e.g. third sector, where there is more experience of this situation). Creating a strong IAPT programme changes the roles for others in adjacent services and teams; where these teams cannot gain

negotiated clarity about their roles, inter-professional tension is likely to be exacerbated.

Working collaboratively with non-NHS organisations and achieving genuine partnership with service users are clearly difficult challenges. There are not enough drivers for these changes to be fully realised, and initial efforts are easily overwhelmed by the NHS having to focus on the job in hand and reverting to traditional modes of working.

8.1.3 The patient experience

Discussing with patients their experience of the IAPT service showed the importance of the first contact with their GP and with the IAPT service in helping to identify the problem, provide hope and a way forward. This was particularly helpful when people had a sense of control and choice and were seen quickly.   Self-referral was often associated with feeling greater self-confidence and hope. However, some patients experienced little or no choice in either referral or treatment options and information that could have helped in decision-making was often not available.

The best experience for patients in terms of guided self-help interventions was characterised by good communication and working with responsive, flexible, and respectful psychological wellbeing practitioners in a structured format tailored to their needs. Negative experience was reported when the practitioner was seen as impersonal, self-help booklets were not pitched at the right level and although there were patients who liked the freedom of telephone contacts and the computerized packages, many found them problematic. Careful introduction, some one-to-one sessions and personal support helped improve the value of telephone/computer working.

Cognitive behaviour therapy was generally valued, but was often thought to be too short.   A lack of continuity (due to staff turnover) or follow-up was problematic for patients.

Feedback from service users commenting on these findings suggest a concern that patients, often feeling vulnerable and lacking in self-confidence, may acquiesce to whatever is offered without a genuine choice of intervention being available. For example, people agreeing to telephone delivery whilst preferring or needing face-to-face contact.   This highlights the importance of genuine choice and consent rather than assuming consent from passive acquiescence in this model of service delivery.

These findings raise the issue of why some clients felt so little real choice, or why they felt the therapist wasn't interested.   These are examples of a failure in establishing and maintaining the therapeutic alliance, which is linked both to patient characteristics and the therapist's skill. Whilst

failures of therapeutic alliance will be found in any psychological therapy service under this degree of scrutiny, it raises the concern that brief interventions offered by relatively inexperienced therapists may, for some patients, offer a poor quality experience. This is not only a concern in itself, but has the potential to deter people from seeking and benefiting from a further, more specialist intervention. For this reason it is important for services specifically to seek feedback on negative impacts of therapy or unsatisfactory patient experience, and to learn from these, rather than rest content with meeting a target for overall levels of good outcomes and client satisfaction.

8.1.4 Service outcomes and costs

In terms of outcomes, both services fell only marginally short of the 50% recovery rate set by the Department of Health as the target for those receiving a minimum treatment of two or more contacts, when a simple index of caseness was used, although somewhat lower recovery rates were found using a more stringent criterion or when the cut-off score of 8 was used for the GAD-7. The target rate of recovery was exceeded when considering those patients who completed their individually agreed treatment plans, where the rates for Doncaster and Newham were 59% and 71% respectively on the least stringent criterion and 50% and 56% respectively on the most stringent.

Helping people return to work and keeping employed people in work, were key aims of the IAPT programme.   Across both sites, there were increases in the proportion of individuals working full time and in Doncaster, a decrease in numbers registered unemployed. In Newham there was an increase in the proportion of people on Incapacity Benefit who described themselves as actively seeking work. However, these changes were in line with benefit off-flow rates for the same period across DWP datasets in these areas, and the cohort study comparison found no differences in changes to employment status between IAPT and comparators. There was evidence from both the cohort study and the general practice datasets that people accessing the IAPT service were taking less time off sick from their jobs than their comparators.

We also investigated how patient outcomes from Doncaster and Newham Demonstration sites compare with those obtained by pre-IAPT primary care services in the UK, and with results reported in research trials, by using a system of benchmarks from archived datasets. These analyses suggest that in terms of recovery rates, patients receiving the minimum dosage fell within the confidence intervals of the benchmark and the recovery rates for

patients completing their planned interventions lay at the upper end of the confidence intervals. That is, the new services delivered a service of equivalent effectiveness, despite being newly-established and delivered by relatively inexperienced practitioners, and this is a considerable achievement.

Comparisons with results reported in research trials showed that the magnitude of the therapy effect for the 'minimum intervention' sample was lower than the benchmark for both depression (1.49) and anxiety (1.42), but for patients completing their planned treatment, the effect size for anxiety approached that for trials while for depression it abutted the lower confidence interval for the research trials. This finding is consistent with a generally observed phenomenon of effects in research trials tending to be rather better than those obtained in routine health service delivery. Many reasons could be and have been suggested for this, for example greater homogeneity and less complex co-morbidity in patients in research samples, selection of patients to those willing to consent to randomisation and the use of more experienced and better supervised therapists.

In addition to studying the clinical outcomes of those accessing the services, we also wished to study the outcomes and service costs for people in primary care eligible for IAPT in the demonstration sites, whether or not they accessed the service, and in comparison sites where IAPT was not available. Unfortunately the pace of the IAPT roll out made this comparison less valid in Newham, where both planned comparator sites acquired IAPT funding during the evaluation period, although this was not the case for Doncaster. Patients were identified by General Practices in Doncaster, Wakefield, Barnsley, Newham and City & Hackney and invited to return self-completed postal questionnaires containing measures of wellbeing, depression, anxiety, health status, employment status and service usage. Return rates for the postal questionnaires were disappointing at 13.6% for the IAPT sample and 10.1% for the comparison sample, with the implication that selection biases have influenced the results. On the other hand these percentages are not grossly different from the self-selected proportion of patients from a given population who agree to enter a treatment trial, and the subsequent randomisation within that selected group does not reduce the external validity problem of how to generalise from results to the whole population.

Outcomes and service use data were collected from 504 patients (self-report) at baseline and at two follow-up points, 4 months and 8 months. The IAPT-site cohorts were generally well matched to their comparison cohorts with no statistically significant differences between them at baseline on socio-demography and clinical measures, although Doncaster respondents reported themselves to be less happy than their comparators

and there were proportionally more men in the Newham sample compared with City and Hackney. Within the IAPT cohort, neither gender nor ethnicity was associated with referral to the service, but younger people were more likely to have accessed the service. At 4 month follow up, 82.5% of the total cohort returned questionnaires, with those who dropped out of the study reporting at baseline more psychological distress and being happy less of the time than those who were retained. At four month follow up, the IAPT cohort and the comparison cohort had improved on all the patient-reported outcome measures; general health, psychological health and wellbeing. Differences between the IAPT cohort as a whole (irrespective of access) and the comparison cohort were confined to those in the IAPT cohort in Doncaster reporting a significant improvement in frequency of happiness but otherwise the two cohorts had similar degree of improvement. Those who accessed the IAPT service were happier and had better SF-6D scores at four months than those who had not accessed the service and those in comparison sites, but after controlling for baseline differences in age and gender, only the difference in happiness remained. At eight months there were no statistically significant differences between the cohorts.

Resource use did not change significantly for most of the mental health services that were reported across the IAPT sites and their matched comparator sites, but GP consultations and other health service use in Doncaster reduced more than in the comparison sites.

Whilst at first glance these results suggest that having IAPT available within a local area made some difference to GP use but little difference to the four or eight month clinical outcome, it is impossible to draw this conclusion with any confidence from these data alone. The low response rate to the postal questionnaire suggests a selection bias meaning that these results may not be generalised to the whole IAPT population. One reason to believe that the sample is not representative is that the clinical outcomes for those who accessed IAPT within the cohort are significantly poorer than the outcomes for the total population of people using service.

This could mean that those who had a poor outcome from IAPT were more likely to return the questionnaire. On the other hand, it is commonly found that people have a response set towards a more favourable response to a questionnaire completed within the context of service receipt compared with an independent postal questionnaire. Where measures are completed frequently as part of a therapeutic relationship, they inevitably form a communication channel between the client and the therapist, and a more positive message is conveyed. In this sense, measurement independent of the therapist is considered more accurate.

Perhaps the key issue is power – do we have sufficient confidence that the lack of difference is real, given the imprecision of the estimate?

Estimation of cost effectiveness in terms of QALY gain for the incremental costs of IAPT over the comparison site was only meaningful when comparing Doncaster with Barnsley and Wakefield. The cohort sample in Newham was too small, and in any case, the comparator site became an IAPT site during the study. The NICE threshold of £30,000 per QALY, below which treatments are considered cost effective, was reached for IAPT when EQ-5D values were used, predicted from the SF-6D based on a regression model estimated on another data set. Using this method for everyone in the Doncaster cohort the incremental cost effectiveness ratio was £20,230 per QALY gained, and for the people in the cohort seen in the IAPT service, £25,950 per QALY. Using national costs rather than IAPT actual costs in the estimate further improved the cost effectiveness estimate at £16,714 and £15,692 per QALY respectively.

It is of interest to compare our cost effectiveness results with the 2007 paper by Layard et al [93], which made an estimation of likely benefits based on a number of assumptions. They assumed face to face CBT rather than the range of interventions being provided by the IAPT services, with a retention rate of 80% and a recovery rate of 60%, persisting for 2 and 5 years. Their NHS analysis assumed a QALY gain of 0.2. In practice, some of these assumptions were optimistic as we did not find this degree of QALY gain and recovery rates were lower. On the basis of these assumptions it is possible to calculate an incremental cost per QALY of £3,630 compared to ours of £25,950, using EQ-5D. They also estimate savings from productivity using assumptions about employment changes based on observational studies, which we did not find. This may be partly due to the small numbers and insufficient length of follow up, so it is not possible to compare our findings on employment directly, but given their assumption of an improvement of 14 percentage points, it is perhaps surprising we did not observe this.

8.1.5 Impact of IAPT at the general practice level

The potential benefits of IAPT at the level of General Practice include reduction of prescriptions for antidepressant medication and referral for unexplained medical symptoms. To investigate this, and to examine the extent of access to IAPT from a general practice perspective, routinely collected patient data from 20 General Practices with access to IAPT services in Doncaster and Newham were analysed. New techniques developed by the St George's Health Informatics Team enabled linkage of GP data with data held by the IAPT services at the individual patient level and with use of secondary care services, without disclosure of patient

identity to the research team. (Data on use of secondary mental health services were not available.) Patients identified as having been referred to IAPT were compared with control patients who were identified as having received a diagnosis of a common MH disorder after 1st April 2007, matched for age, gender, and practice.

Since 1st April 2007, the 12 month prevalence of common mental disorder was 10% for the adult population of these practices. Overall only a small proportion (6.3%) of patients with a recorded diagnosis of a common mental health disorder was referred to the IAPT services by the participating practices during the year from 1st October 2007. On the other hand, a substantial proportion of people referred to IAPT (31%) did not have a recorded diagnosis of a common mental health problem. Some of these had other diagnoses, such as difficulty sleeping or alcohol misuse. It is possible that those who had no recorded diagnosis may have accessed IAPT through an alternative service or self-referral.

Associations were found between referral to IAPT and higher levels of psychotropic prescriptions and sickness certification but lower levels of secondary physical health service usage. In order to understand the timing and direction of these associations we compared these variables before and after referral to the IAPT service. This revealed that the level of overall antidepressant prescribing was not reduced by access to the IAPT service, although the proportion of people receiving medication did reduce in the IAPT patients compared to those who were not referred. Sickness certification reduced among the IAPT referred patients after referral but increased slightly among the controls during the same time period, and this difference was statistically significant. This is in line with other findings in relation to the IAPT services [36] and suggests that patients with higher levels of sickness certification are more likely to be referred to IAPT and these levels reduce as a result. Findings on the use of secondary care services were mixed, with a modest reduction in A&E attendances and a trend towards reduced inpatient days but a small increase in outpatient attendances for the IAPT referrals compared with their controls.

8.1.6 Conclusions

Overall, our findings suggest that IAPT referral is being appropriately targeted on people with a greater severity of problem, sickness certification and use of medication, and although it is not reducing antidepressant prescribing overall, it seems to be reducing sickness certification and may lead to a reduction in the use of Accident and Emergency attendances. These potential benefits at the practice level are diluted by the small proportion of people with common mental health problems who are referred.

The IAPT programme represents the biggest central funding commitment to psychological therapies in the history of the NHS and its first systematic attempt to respond to the chronic under-provision of services in this field. Results from the demonstration sites show that both services were successfully established and offered good access to collaborative care for people with common mental health problems. Results met Department of Health expectations and were equivalent to psychological therapies delivered by other primary care practitioners, with evidence of reduction in sickness certification and possibly in the use of some secondary health services. As Gilbody and his colleagues point out [94], in a population-level strategy to improve the management of depression in a greater number of patients, even a modest effect can reduce the overall burden of illness associated with depression.

Limitations of the evaluation

There were many difficulties in mounting an evaluation of a new model of services in this field and limitations in the design and implementation of this evaluation[52]. In this section we shall list some of the limitations of the study, the main problems which affect interpretation of our findings, and make some recommendations for how these difficulties could be avoided in future.

Any use of routinely collected data for research purposes brings the danger of imprecision or misinterpretation, the former because not all data fields are completed reliably for all patients and the latter because the reasons practitioners complete fields in a certain way may be idiosyncratic and not apparent to an outside observer. For example, 'diagnosis' was recorded very differently between the two sites and there is likely imprecision here. We also had the difficulty for one site of some records of earlier episodes being overwritten for a new episode, making the resultant 'hybrid' record undecipherable for research purposes. Results presented in chapters 2, 5 and 7 are subject to this caveat. In chapter 5, when interpreting results based on effect sizes, it needs to be kept in mind that these are simply point estimates of the population effect and will vary from study to study. There are difficulties of potential selection bias when data are only available from a self-selected group of practices or of patients. Results presented in

---

[52] Some of these issues for policy research have been discussed by Professor Chris Salisbury at the University of Bristol in a 2010 report commissioned by the Department of Health on evaluation projects arising from the NHS white paper "Our Health, Our Care, Our Say".

chapters 6 and 7 are subject to this caveat. In relation to chapter 6, the poor response to our questionnaire improved considerably when we offered a small incentive in the form of a shopping voucher for each questionnaire returned, and it would have been preferable to have offered this from the outset and to have included the costs of this in the funding structure for the project. Qualitative results are also dependent on the integrity of the process for generating them, and our intention to capture a purposive sample for Chapter 4 (patient experience) was frustrated by non-response for some groups more than others, which results in another form of selection bias. The research team included a service user researcher, who in addition to generic research skills gave a user perspective on a range of issues, which resulted in, for example, changes to recruitment procedures. However, there could have been greater service user involvement throughout the project; a panel of users were invited to comment on a draft of this report but only a few participated. It would have been preferable to engage members of this panel earlier and to have had structured opportunities for their input throughout.

There were three main difficulties in designing and implementing this evaluation which may have broader relevance for other research of this type; the fast-moving nature of the policy under evaluation compared with the slow-moving process of commissioning and initiating the evaluation, the delays introduced by research governance processes and the difficulty in engaging General Practice in cooperating with the research.

Following a tendering process that lasted almost a year, we were contracted to undertake this evaluation late in 2006, whereas the IAPT services were being commissioned a year earlier. After appointing staff, seeking research governance approvals and gaining collaboration from General Practices, we did not start data collection until May 2007. By this time, the Department of Health had issued a press release entitled "Hewitt hails talking therapy pilots a success and announces more to follow". This could be seen to undermine a robust evaluation, especially since a randomised approach to testing the innovation had already been ruled out by the tender specification and then two of our comparator sites were selected for the IAPT roll out. This points to a tension between the needs of the Department for very quick results in order to grasp an opportunity for policy implementation and the needs of a formal, independent and systematic research evaluation to be subject to competitive national tendering and research governance.

We recommend that in future the mismatch between the timescales for useful policy results and for NIHR-commissioned independent research be addressed for programmes of this type. Possibilities include a) the use of the DH policy research units for shorter-scale early evaluation which

remains independent of the programme board and b) incorporating an element of cluster randomisation into the roll-out programme to allow a robust estimate of the incremental benefits of the programme. The 'stepped wedge' designs mentioned in the phase 2 MRC Guidance on developing and evaluating complex interventions seem particularly appropriate here.

There is no doubt that burgeoning research governance processes have become extremely time-consuming and demanding for researchers and were possibly disproportionate in this case. For example, when seeking permission to approach General Practices to send out postal questionnaires in one area, we had to argue with the PCT research office that it was unnecessary for all 15 members of our research team to have honorary NHS contracts with the Trust in question, including occupational health clearance and criminal record bureau checks.

Finally, research of this kind is fundamentally dependent on co-operation from primary care practitioners, and General Practices are often reluctant to co-operate with research, even when commissioned by the Department of Health or NIHR. Whilst the development of the Primary Care Research Network was seen as a way to improve this problem, in practice we did not find this helped us recruit practices and some more systemic solution to ensuring NHS co-operation with NHS-commissioned research seems important.

Lessons for the future development of IAPT

As IAPT continues to develop across England, with implications for services in other parts of the UK, here we give some advice for the continuing implementation of this worthwhile service innovation.

Overall there was little evidence of major inequities in access to IAPT with the clear exception of age discrimination, with older adults under-represented in referrals to these services, although there is no evidence that they would not benefit. This suggests that there may be some misgivings on the part of GPs and other referrers about the appropriateness of these services for older people, or attitudinal barriers for older adults accessing services. An early emphasis in making the economic case for IAPT was on the potential of these therapies to allow adults of working age on welfare benefits due to their ill health to return to paid employment and to reduce the risk of long term sickness absence or disability for current employees. Although neither service included an upper age limit in their referral criteria, this may have influenced the general perception of the

target groups for the services. One implication of this is the need for services explicitly to include the older population in their outreach activities, including community liaison and local social marketing to these groups. There is also a dearth of research on barriers to access to psychological therapies for older adults with common mental health problems who could benefit from collaborative care.

Although access to collaborative care was increased through IAPT, it remains the case that only a small proportion of people with common mental disorders within primary care were referred, and of these, more than half did not go on to receive an intervention. Given the disparity between the numbers of people with depression and anxiety disorders seen in General Practice and the capacity of even a low-intensity IAPT service, some targeting of the referrals is inevitable. Whereas it seemed from the GP datasets that people with more severe or acute problems are being appropriately referred, local examination of the reasons people do not go on to receive an intervention is justified. It may be possible to refine the referral practices of GPs to increase this proportion, or it may be accepted that there will be a group who do not engage but for whom the advice and signposting function of the IAPT services is both necessary and valued.

The distinctive aspects of IAPT services are important in giving them a unique identity, and this special identity was found helpful in establishing the services. As IAPT services develop more generally and are commissioned alongside other services, sustaining the 'IAPT vision' of community-based, de-stigmatising and accessible care will be an important task in ensuring they do not revert to a more provider-focussed rather than user-centred model.

Partnership working was perhaps the most difficult challenge faced by the services (and given the magnitude of the other challenges, this is a significant point). This should be an important priority for developing services across the country. Three types of partnerships are involved; with other providers within the care pathway (e.g. GPs and CMHTs), with non-NHS agencies such as employers' occupational health services, third sector organisations, jobcentres, condition management programmes, and with service users. Given the salience of this issue, we recommend that services dedicate staff time to the organisational development needs of these partnerships, including some of the helpful interventions outlined in Chapter 3.

In terms of patient acceptability, the need for sensitive handling of choice and consent issues is apparent, with awareness that for some people a telephone-administered 'self-help' approach will cause difficulties. It could be argued that in organising and delivering stepped care there is an intrinsic

tension between fast access, therapy effectiveness and genuine choice: perhaps this is analogous to the business aphorism 'Good. Fast. Cheap. Choose any two.' However, as is already happening in a number of areas, easy referral to alternative services increases choice for individuals who find it hard to engage with IAPT, and supports the importance of the signposting function of the IAPT process. At the other end of the process, ensuring appropriate follow up and continuity of care after the IAPT intervention has been made is an area that services should take care not to neglect.

Some service users found repeated measures intrusive resulting in lower satisfaction with the service as well as poor validity of measurement. Staff training should enhance the ability of staff to handle the measurement requirement in ways that are sensitive to the needs of patients and do not damage the therapeutic alliance.

Given the costs of repeated measurement, it is important to use the resulting datasets to the maximum, not only for central reporting and evaluation but for local service quality improvement. Services can sometimes struggle to gain the maximum local benefit from the large quantities of data they collect. Simple quality improvement methods can be used where staff are encouraged to review examples of long waiting time, poor outcomes or drop out in order to generate ideas on remediable reasons for these.

Patients spoke of the need for good communication and working with responsive, flexible, and respectful psychological wellbeing practitioners in a structured format tailored to their needs. This is the kernel of providing a high quality service; the personalisation of care and the training and support that practitioners need to be able to provide it.

Recommendations

8.1.7 Recommendations for NHS practice

1. Our findings endorse the current policy of including a wider range of evidence-based therapeutic modalities within the IAPT programme, both in terms of benchmarked outcomes and patient choice.

2. IAPT services should emphasise to referrers that they welcome referrals of older adults. Community outreach activities, including community liaison and local social marketing, should be targeted to groups of older adults.

3. We recommend that NHS data collection systems a) retain a record of closed episodes linked by patient ID rather than overwrite old records b) include NHS number.

4. It is important to use routinely collected data for service quality improvement locally as well as reporting them centrally. We recommend local audit and follow up of cases of a) clinically significant and statistically reliable deterioration, b) variance outliers in waiting times (latency), and c) drop out following early engagement. Similarly, feedback should be sought on negative patient experience and remediable reasons for this explored.

5. As part of partnership working with referrers, there should be opportunity for joint review of those cases where referral was followed by a failure to engage, to explore the reasons for lack of engagement and ways in which the proportion of those referred who complete an intervention can be increased.

6. Services should continue to emphasise choice and to offer genuine choice of interventions at the same service step following initial assessment. This implies actively promoting the 'signposting' function of assessment.

7. We recommend services review their provision of appropriate follow up and aftercare subsequent to the IAPT intervention.

8. We recommend that further methods of ensuring NHS collaboration with NHS-commissioned research be explored.

9. The lack of NHS Number in the minimum dataset for the IAPT clinic makes audit and analysis where data-linkage is needed much more challenging; we recommend this should be included in NHS IAPT databases.

10. Meaningful service user involvement in service provision is required to enable user-centred service design and to address acceptability issues.

11. We recommend that services dedicate staff time to the organisational development needs of partnerships with other NHS providers, non-NHS providers and service users. This requires commitment from a high level in the organisation, and mechanisms for front-line staff to have a voice in resolving partnership process issues.

8.1.8 Recommendations for research

1. Two types of research on the impact of IAPT on secondary mental health services would be valuable as our study was not able to address this. First, a replication of the General Practice data analysis on the impact of IAPT referral vs. non-referral is needed, covering referral to and use of secondary mental health care. Second, little is known about the impact of IAPT locally on secondary and tertiary mental health services from an organisational perspective.

2. Research on psychological service outcomes should include measures that go beyond symptomatology and that service users find relevant, for example, better assessment of functioning and quality of life.

3. A study of the difference in scores between therapist-administered and researcher-administered measures in IAPT service users would be relatively simple to undertake and extremely useful in future estimation of true effects of treatment.

4. Given the difficulty in recruitment experienced in this and other similar studies, research is recommended on the most effective ways of engaging patients with depression and other mental health problems in research participation.

5. The routine datasets collected as part of the IAPT roll-out represent an important and costly NHS resource, which should be used to the maximum to support a wide range of research. We recommend that to gain greatest benefit from this investment, research access to these anonymised datasets should be available to bona fide research groups where protocols have been independently scientifically reviewed.

6. Further in-depth exploration of service user views, beyond the scope of that reported here, would be of great value; in particular, we did not compare service user experience of IAPT with that of other psychological services.

7. Evaluation of IAPT outcomes in relation to new service and trials benchmarks should be undertaken, particularly as more trials on self-help and computer-supported interventions become available.

References

[1] Layard R. 2004. *Mental health: Britain's biggest social problem? Paper presented at the No.10 Strategy Unit Seminar on Mental Health.*

[2] National Collaborating Centre for Mental Health. 2009. *Depression: The treatment and management of depression in adults.* London, National Institute for Health and Clinical Excellence.

[3] National Collaborating Centre for Mental Health. 2009. *Depression: The treatment and management of depression in adults.* London, National Institute for Health and Clinical Excellence.

[4] National Collaborating Centre for Mental Health. 2004 *Anxiety: Clinical guideline.* London: National Institute for Health and Clinical Excellence.

[5] National Collaborating Centre for Mental Health. 2009. *Depression: The treatment and management of depression in adults.* London, National Institute for Health and Clinical Excellence.

[6] Department of Health. 2001. *Treatment choice in psychological therapies and counselling.* London, HMSO.

[7] Singh S. 2005. *CBT service report to Newham Primary Care Trust board.*

[8] Newham IAPT development team. 2005. *Newham project initiation document v7, 12th November.*

[9] Evans C, Connell J, Barkham M, Marshall C, Mellor-Clark J. 2003; Practice-based evidence: Benchmarking NHS primary care counselling services at national and local levels. *Clinical Psychology and Psychotherapy* 10:374-388.

[10] Greenhalgh T, Robert G, Bate P, Kyriakidou O, Macfarlane F. 2004. *How to spread good ideas: A systematic review of the literature on diffusion, dissemination and sustainability of innovations in health service delivery and organisation.*   National Co-ordinating Centre for NHS Service Delivery

and Development.

[11] Pettigrew AM. 1990; Longitudinal field research on change: Theory and practice. *Organization Science* 1(3):267-292.

[12] Pettigrew AM, Woodman RW, Cameron KS. 2001; Studying organizational change and development: Challenges for future research. *Academy of Management Journal* 44(4):697-713.

[13] Pawson R, Tilley N. 1997. *Realistic Evaluation.* London, Sage Publications.

[14] Retolaza A, Grandes G. 2003; Expectations and satisfaction in mental health center users. *Actas Espanolas de Psiquiatria* 31(4):171-176.

[15] Attkisson CC, Zwick R. 1982; The client satisfaction questionnaire: Psychometric properties and correlations with service utilization and psychotherapy outcome. *Evaluation and Program Planning* 5(3):233-237.

[16] Larsen DL, Attkisson CC, Hargreaves WA, Nguyen TD. 1979; Assessment of client/patient satisfaction: Development of a general scale. *Evaluation and Program Planning* 2:197-207.

[17] Braun V, Clarke V. 2006; Using thematic analysis in psychology. *Qualitative Research in Psychology* 3:77-101.

[18] Lewis J. 2003. Design issues. In: J.Ritchie, J.Lewis, editors. *Qualitative Research Practice: A Guide for Social Science Students and Researchers.* London: Sage Publications; 47-76.

[19] Smith JA, Flowers P, Larkin M. 2009 *Interpretive Phenomenological Analysis: Theory, Method and Research.* London: Sage.

[20] Kroenke K, Spitzer RL, Williams JBW. 2001; The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine* 16:606-613.

[21] Spitzer RL, Kroenke K, Williams JBW, Löwe BA. 2006; A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives Internal Medicine* 22:1092-1097.

[22] Merrill KA, Tolbert VE, Wade WA. 2003; Effectiveness of cognitive therapy for depression in a community mental health center: A benchmarking study. *Journal of Consulting and Clinical Psychology* 71:404-409.

[23] Wade WA, Treat TA, Stuart GL. 1998; Transporting an empirically supported treatment for panic disorder to a service clinic setting: A benchmarking strategy. *Journal of Consulting and Clinical Psychology* 66:231-239.

[24] Mullin T, Barkham M, Mothersole G, Bewick BM, Kinder A. 2006; Recovery and improvement benchmarks in routine primary care mental health settings. *Counselling and Psychotherapy Research* 6:68-80.

[25] Paley G, Cahill J, Barkham M, Shapiro D, Jones J, Patrick S. 2008; The effectiveness of psychodynamic-interpersonal psychotherapy in routine clinical practice: A benchmarking comparison. *Psychology and Psychotherapy: Theory, Research and Practice* 85:157-175.

[26] Minami T, Wampold BE, Serlin RC, Hamilton EG, Brown JS, Kircher JC. 2008; Benchmarking the effectiveness of psychotherapy treatment for adult depression in a managed care environment: A preliminary study. *Journal of Consulting and Clinical Psychology* 76:116-124.

[27] Barkham M, Margison F, Leach C, Lucock M, Mellor-Clark J, Evans C et al. 2001; Service profiling and outcomes benchmarking using the CORE-OM: Towards practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology* 69:184-196.

[28] Department of Health. 2010 *Realising the benefits: IAPT at full roll out.* Department of Health.

[29] Richards DA, Suckling R. 2009; Improving access to psychological therapies: Phase IV prospective cohort study. *British Journal of Clinical Psychology*.

[30] Barkham M, Gilbert N, Connell J, Marshall C, Twigg E. 2005; Suitability and utility of the CORE-OM and CORE-A for assessing severity of presenting problems in psychological therapy services based in primary and secondary care settings. *Br J Psychiatry* 186:239-246.

[31] Evans C, Mellor-Clark J, Margison F, Barkham M, McGrath G, Connell J et al. 2000; Clinical outcomes in routine evaluation: The CORE-OM. *Journal of Mental Health* 9:247-255.

[32] Evans C, Connell J, Barkham M, Margison F, Mellor-Clark J, McGrath G et al. 2002; Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE-OM. *Br J Psychiatry* 180:51-60.

[33] Mellor-Clark J, Barkham M. 2006. The CORE system: Developing and delivering practice0-based evidence through quality evaluation. In:

Feltham C, Horton I, editors. *Handbook of Counselling and Psychotherapy.* 2 ed. London: Sage Publications; 207-224.

[34] Department of Health. 2010 *Improving Access to Psychological Therapies (IAPT) outcomes toolkit 2008/9.* Author.

[35] Beck A, Steer R, Garbin G. 1988; Psychometric properties of the Beck Depression Inventory: Twenty five years of evaluation. *Journal of Consulting and Clinical Psychology* 56:893-897.

[36] Clark DM, Layard R, Smithies R, Richards DA, Suckling R, Wright B. 2009; Improving Access to Psychological Therapy: Initial evaluation of two UK demonstration sites. *Behaviour Research and Therapy* 47(11):910-920.

[37] Evans C, Margison F, Barkham M. 1998; The contribution of reliable and clinically significant change methods to evidence based mental health. *Evidence Based Mental Health* 1:70-72.

[38] Jacobson NS, Traux P. 1991; Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology* 59:12-19.

[39] Connell J, Barkham M, Stiles WB, Twigg E, Singleton N, Evans O et al. 2007; Distribution of CORE-OM scores in a general population, clinical cut-off points and comparison with the CIS-R. *Br J Psychiatry* 190:69-74.

[40] Cahill J, Barkham M, Stiles W. 2010; Systematic review of practice-based research on psychological therapies in routine clinic settings. *British Journal of Clinical Psychology* 49:421-453.

[41] Gibbard I, Hanley T. 2008; A five year evaluation of the effectiveness of person centred counselling in routine clinical practice in primary care. *Counselling and Psychotherapy Research* 8:215-222.

[42] Gilbert N, Barkham M, Richards A, Cameron I. 2005; The effectiveness of a primary care mental health service delivering brief psychological interventions: A benchmarking study using the CORE system. *Primary Care Mental Health* 3:241-251.

[43] Mellor-Clark J, Connell J, Barkham M, Cummins P. 2001; Counselling outcomes in primary health care: A CORE system data profile. *European Journal of Psychotherapy* 4:65-86.

[44] Shepherd M, Ashworth M, Evans C, Robinson SI, Rendall M, Ward S. 2005; What factors are associated with improvement after brief psychological interventions in primary care? Issues arising from using routine outcome

---

measurement to inform clinical practice. *Counselling and Psychotherapy Research* 5:273-280.

[45]   Conway S, Audin K, Barkham M, Mellor-Clark J, Russell S. 2003; Practice-based evidence for a brief time-intensive multi-modal therapy guided by group-analytic principles and method. *Group Analysis* 36:413-435.

[46]   Gordon K, Graham C. 1996; The impact of primary care counselling on psychiatric symptoms. *Journal of Mental Health* 5:515-523.

[47]   Stiles W, Barkham M, Mellor-Clark J, Connell J. 2008; Effectiveness of cognitive behavioural, person-centred, and psychodynamic therapies in UK primary-care routine practice: Replication in a larger sample. *Psychological Medicine* 38:677-688.

[48]   Westbrook D, Kirk J. 2005; The clinical effectiveness of cognitive behaviour therapy:Outcome for a large sample of adults treated in routine practice. *Behaviour Research and Therapy* 43:1243-1261.

[49]   National Collaborating Centre for Mental Health. 2009. *Depression: The treatment and management of depression in adults.* London, National Institute for Health and Clinical Excellence.

[50]   Furukawa T, Watanabe N, Churchill R. 2007; Combined psychotherapy plus antidepressants for panic disorder with or without agoraphobia. *Cochrane Database of Systematic Reviews* 1.

[51]   Hunot V, Churchill R, Teixeira V, Siva de Lima M. 2007; Psychological therapies for Generalised Anxiety Disorder. *Cochrane Database of Systematic Reviews* 1.

[52]   Thomson A, Page L. 2007; Psychotherapies for hypochondriasis. *Cochrane Database of Systematic Reviews* 4.

[53]   Clark D. 2006. *Effective psychological treatments for anxiety disorders: A report for the Department of Health in support of the submission to the Comprehensive Spending Review.*

[54]   Bisson J, Andrew M. 2007. *Psychological treatment of post-traumatic stress disorder (PTSD).* Cochrane Database of Systematic Reviews 3.

[55]   Ehlers A, Clark D, Hackmann A, McManus F, Fennell M, Herbert C. 2003; A randomized controlled trial of cognitive therapy, a self help booklet, and

repeated assessments as early interventions for posttraumatic stress disorder. *Archives of General Psychiatry* 60:1024-1032.

[56] Brown J, Elliott S, Boardman J, Ferns J, Morrison J. 2004; Meeting the unmet need for depression services with psycho-educational self-confidence workshops: preliminary report. *Br J Psychiatry* 185:511-515.

[57] Lovell K, Bower P, Richards D, Barkham M, Sibbald B, Roberts C et al. 2008; Developing guided self-help for depression using the Medical Research Council complex interventions framework: A description of the modelling phase and results of an exploratory randomised controlled trial. *Bmc Psychiatry* 8.

[58] Kessler D, Lewis G, Kaur S, Wiles N, King M, Weich S. 2009; Therapist-delivered internet psychotherapy for depression in primary care: a randomised controlled trial. *The Lancet* 374:628-634.

[59] Proudfoot J, Ryden C, Everitt B, Shapiro D, Goldberg D, Mann A. 2004; Clinical efficacy of computerised cognitive-behavioural therapy for anxiety and depression in primary care: randomised controlled trial. *Br J Psychiatry* 185:46-54.

[60] Purves D, Bennett M, Wellman N. 2009; An open trial in the NHS of Blues Begone: A new home based computerized CBT program. *Behavioural and Cognitive Psychotherapy* 37:541-551.

[61] Salkovskis P, Rimes K, Stephenson D, Sacks G, Scott J. 2006; A randomized controlled trial of the use of self-help materials in addition to standard general practice treatment of depression compared to standard treatment alone. *Psychological Medicine* 36:325-333.

[62] Ward E, King M, Lloyd M, Bower P, Sibbald B, Farrelly S. 2000; Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy, and usual general practitioner care for patients with depression. I: Clinical effectiveness. *British Medical Journal* 321:1383-1388.

[63] Barrowclough C, King P, Russell E. 2001; A randomized trial of effectiveness of cognitive behavioural therapy and supportive counselling for anxiety disorders in older adults. *Journal of Consulting and Clinical Psychology* 69:756-762.

[64] Clark D, Ehlers A, Hackmann A, McManus F, Fennell M, Gret N. 2006; Cognitive therapy versus exposure and applied relaxation in social phobia: A randomized controlled trial. *Journal of Consulting and Clinical Psychology* 74:568-578.

[65] Clark D, Ehlers A, McManus F, Hackmann A, Fennell M, Campbell H. 2003; Cognitive therapy versus fluoxetine in generalized social phobia: a randomized placebo controlled trial. *Journal of Consulting and Clinical Psychology* 71:1058-1067.

[66] Clark D, Salkovskis P, Hackmann A, Wells A, Fennell M, Ludgate J. 1998; Two psychological treatments for hypochondriasis: A randomised controlled trial. *Br J Psychiatry* 173:218-225.

[67] Clark D, Salkovskis P, Hackmann A, Wells A, Ludgate J, Gelder M. 1999; Brief cognitive therapy for panic disorder: a randomized controlled trial. *Journal of Consulting and Clinical Psychology* 67:583-589.

[68] Warwick H, Clark D, Cobb A. 1996; A controlled trial of cognitive-behavioural treatment of hypochondriasis. *Br J Psychiatry* 169:189-195.

[69] Statacorp. 2001 *Statistical Software: Release 7.0.* College Station, TX: Stata Corporation.

[70] Barkham M, Stiles W, Connell J, Twigg E, Leach C, Lucock M. 2008; Effects of psychological therapies in randomized trials and practice-based studies. *British Journal of Clinical Psychology* 47:397-415.

[71] Brazier J, Roberts J, Deverill M. 2002; The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics* 21(2):271-292.

[72] National Institute for Clinical Excellence (NICE). 2004 *Guide to the methods of technology appraisal.* NICE London.

[73] Curtis L. 2009. *The Unit costs of Health and Social Care.* Personal Social Services Research Unit.
Ref Type: Generic

[74] Department of Health. 2010. *NHS reference costs 2008-2009.*
Ref Type: Online Source

[75] Munro J, Nicholl J, O'Cathain A, Knowles E, Morgan A. 2001 *Evaluation of NHS Direct first wave sites: Final report of the phase 1 research.* Medical Care Research Unit. University of Sheffield.

[76] Office of National Statistics. 2009. *Annual Survey of Hours and Earnings (ASHE) 2009.*

[77] Petrou S, Hockley C. 2005; An investigation into the empirical validity of the EQ-5D and SF-6D based on hypothetical preferences in a general population. *Health economics* 14(11):1169-1189.

[78] Office of National Statistics. 2010. *Labour Market Statistics.*

[79] Evans C, Crawford B. 1999; Patient self-reports in pharmacoeconomic studies: their use and impact on study validity. *Pharmacoeconomics* 15(3):241-256.

[80] Petrou S, Murray L, Cooper P, Davidson LL. 2002; The Accuracy of Self-reported Healthcare Resource Utilization in Health Economic Studies. *International Journal of Technology Assessment in Health Care* 18(03):705-710.

[81] Byng R, Gask L. 2009; Improving Access to Psychological Therapies: Implications for mental health care in general practice. *British Journal of General Practice* 59:640-641.

[82] Hague J. 2008; IAPT: What does it mean for primary care? *Healthcare Counselling and Psychotherapy Journal*:11-14.

[83] Layard R. 2006; The case for psychological treatment centres. *British Medical Journal* 332(7548):1030-1032.

[84] Huffmann JC, Pollock MH. 2003; Predicting panic disorder among patients with chest pain: An analysis of the literature. *Psychosomatics* 44:222-236.

[85] Chan T, Kumarapeli P, van Vlymen J, Hasan A, de Lusignan S. 2007 *Improving Access to Psychological Therapies pilots in Doncaster and Newham: Interim report of primary care data.* University of London.

[86] Moore M, Yuen HM, Dunn N, Mullee MA, Maskell J, Kendrick T. 2009; Explaining the rise in antidepressant prescribing: A descriptive study using the general practice research database. *British Medical Journal* 339:399.

[87] McManus S, Meltzer H, Brugha T, Bebbington P, Jenkins R. 2009 *Adult psychiatric morbidity in England, 2007: Results of a household survey.* London: National Centre for Social Research.

[88] Goldberg D, Huxley P. 1980 *Mental Illness in the Community: The Pathway to Psychiatric Care.* London: Tavistock Publications.

[89] Goldberg D, Privett M, Uston B. 1998; The effects of detection and treatment on the outcome of major depression in primary care: A

naturalistic study in 15 cities. *British Journal of General Practice* 48:1840-1844.

[90]   Kessler D, Lloyd K, Lewis G, Pereira Gray D. 1999; Cross sectional study of symptom attribution and recognition of depression and anxiety in primary care. *British Medical Journal* 318:436-440.

[91]   Kendrick T, Dowrick C, McBride A, Howe A, Clarke P, Maisey S et al. 2009; Management of depression in UK general practice in relation to scores on depression severity questionnaires: Analysis of medical record data. *British Medical Journal* 338:750.

[92]   de Lusignan S, van Weel C. 2010; The use of routinely collected computer data for research in primary care: opportunities and challenges. *Family Practitioner* 23(2):253-263.

[93]   Layard R, Clark D, Knapp M, Mayraz G. 2007; Cost-benefit analysis of psychological therapy. *National Institute Economic Review* 202:90.

[94]   Gilbody S, Bower P, Fletcher J, Richards D, Sutton A. 2006; Collaborative care for depression: A cumulative meta-analysis and review of longer-term outcomes. *Archives of Internal Medicine* 166:2314-2321.

Addendum

This document is an output from a research project that was commissioned by the Service Delivery and Organisation (SDO) programme whilst it was managed by the National Coordinating Centre for the Service Delivery and Organisation (NCCSDO) at the London School of Hygiene & Tropical Medicine. The NIHR SDO programme is now managed by the National Institute for Health Research Evaluations, Trials and Studies Coordinating Centre (NETSCC) based at the University of Southampton.

Although NETSCC, SDO has managed the project and conducted the editorial review of this document, we had no involvement in the commissioning, and therefore may not be able to comment on the background of this document. Should you have any queries please contact sdo@southampton.ac.uk.