

Department for Work and Pensions

Working Paper No 45

Understanding the impact of JRRP for people with mental health conditions

Rebecca Taylor and Jane Lewis

A report of research carried out by the Department for Work and Pensions

© Crown Copyright 2008. Published for the Department for Work and Pensions under licence from the Controller of Her Majesty's Stationery Office.

Application for reproduction should be made in writing to The Copyright Unit, Her Majesty's Stationery Office, St Clements House, 2-16 Colegate, Norwich NR3 1BQ.

First Published 2008.

ISBN 978 1 84712 371 8

Views expressed in this report are not necessarily those of the Department for Work and Pensions or any other Government Department.

Contents

Acknowledgements	v
The Authors.....	vi
1 Introduction.....	1
1.1 JRRP and the research objective	1
1.2 The data sources.....	3
1.2.1 <i>Quantitative data sources</i>	3
1.2.2 <i>Qualitative data sources</i>	4
1.3 The approach taken in the reanalysis study	6
2 Descriptive analysis	9
2.1 Who are they?.....	9
2.2 What happened in the pilot?	10
2.2.1 <i>Health-related treatments</i>	11
2.2.2 <i>Work-related treatments</i>	12
2.2.3 <i>Advice</i>	13
2.2.4 <i>Why specific interventions were offered</i>	13
2.2.5 <i>Contact as an intervention</i>	14
2.2.6 <i>Summary</i>	14
3 Explanatory analysis: the hypotheses	15
3.1 H1: despite randomisation, there are significant differences in the characteristics of the MH-IG and MH-CG	15
3.2 H2: the finding reflects the definition of mental health as primary condition and is not replicated with other definitions of mental health	16

3.3	H3: the finding reflects the definition of the primary outcome and is not replicated with other definitions	17
3.3.1	<i>Full-time</i>	17
3.3.2	<i>By week 29/33</i>	17
3.3.3	<i>No break in employment</i>	18
3.3.4	<i>13 continuous weeks</i>	18
3.3.5	<i>Survey non-response</i>	18
3.3.6	<i>Other outcomes</i>	20
3.3.7	<i>Conclusion</i>	20
3.4	H4: the JRRP services were more deficient for the MH-IG than for the non-MH-IG.....	21
3.5	H5: the quality of NHS treatment for mental conditions leaves less scope for JRRP to boost treatment for MH-IG than non-MH-IG	21
3.6	H6: the workplace context leaves less scope for JRRP to boost treatment for MH-IG than non-MH-IG	23
3.7	H7: the focus on returning to the same employer was inappropriate for the MH-IG	24
3.8	H8: the MH-IG experienced less encouragement to return to work before they were fully fit	26
3.9	H9: JRRP differentially encouraged dependency/discouraged proactivity among the MH-IG	27
4	Conclusions	31
	References	33

List of tables

Table 1.1	Return to work rate mental health participants	2
Table 2.1	Comparison of characteristics for MH and non-MH participants.....	10
Table 2.2	Health treatments received	11
Table 2.3	Work treatments received	12
Table 2.4	Advice received.....	13
Table 2.5	Main reasons for offering treatments.....	13
Table 3.1	Return to work rate for mental health participants: different weights	19
Table 3.2	Measures of supportiveness of the workplace	23

Acknowledgements

We would like to offer particular thanks to our colleagues Susan Purdon and John D'Souza for their help and advice with the quantitative analysis, and the valuable comments arising from our presentation of these findings at the Department for Work and Pensions. We are grateful to all the people who have been involved with JRRP over the last few years.

The Authors

Jane Lewis was the Director of NatCen's Qualitative Research Unit until October 2007 and is now Director of Research, Evidence and Evaluation at the National Children's Bureau. At NatCen she co-directed the qualitative research element of the JRRP evaluation, directed qualitative and mixed-methods research studies and carried out methodological work on the role and application of qualitative methods in policy research.

Rebecca Taylor is a Research Director in NatCen's Survey Methods Unit. She worked as a Research Advisor for the evaluation of JRRP. More recently her work has focused on data collection methodology, such as methods to improve response rates and experiments to learn more about non-response from longitudinal surveys.

1 Introduction

1.1 JRRP and the research objective

The purpose of this reanalysis study was to explore an unexpected finding emerging from the evaluation of the Job Retention and Rehabilitation Pilot (JRRP), namely that people with mental health conditions who used the JRRP services had a lower rate of return to work than those who did not use the service.

JRRP was developed by the Department for Work and Pensions (DWP) to test interventions which might improve the return to work rate of people off sick from work. The pilot ran as a randomised control trial for a period of nearly two years from April 2003 in six areas of the UK, operated by four different Providers.

Participation in the pilot was voluntary. Potential participants were accepted on to the pilot if they had been on sickness absence from employment for between six and 26 weeks and passed a screening test designed to exclude people with a high likelihood of returning to work without any intervention. Participants were assigned randomly to one of three intervention groups or to a control group. The three intervention groups were:

- Health intervention group – providing only services to address participants' health conditions. These services included:
 - access to medical tests, consultations, treatment and rehabilitation programmes;
 - various forms of psychological therapy;
 - physiotherapy and hydrotherapy;
 - complementary therapies, including aromatherapy, reflexology, and massage;
 - exercise programmes and facilities, including membership of gyms.
- Workplace intervention group – providing only services to address workplace issues which included:
 - mediation with employers;

- ergonomic and equipment assessments at the workplace;
- occupational therapy;
- advice on welfare rights;
- career advice;
- CV preparation and jobsearch support;
- Combined intervention group – providing both health and workplace services and thus able to offer any of the services listed above.

The services were delivered by a combination of in-house staff and external organisations or individuals, either working in a partnership arrangement with the Provider or commissioned on an ad hoc basis to provide a particular type of support. Common to all Provider organisations however was the allocation of each client to a member of staff who acted as case manager (though not all organisations used this term). The case managers acted as the focal point of contact for clients, gathering information from them, advising them on options, acting as gatekeepers to other services, and in some cases providing actual services themselves. For a more detailed description of the pilot see Stratford et al, 2005b.

The main assessment of the trial, the impact evaluation, did not find any differences in the return to work rate between the intervention groups and the control group for participants as a whole (Purdon et al, 2006). However this was not the case for the one-third of participants that reported a mental health condition as their primary reason for being off work when they entered the pilot. Unexpectedly, the trial showed that people with a mental health condition as their primary condition were significantly less likely to return to work if they were in one of the three intervention groups than if they were in the control group i.e. the service appeared to have a negative impact (Table 1.1). A comparison with all participants shows that the return to work rates in the intervention groups are of a similar magnitude to each other, in other words, the findings were not driven by one particular intervention.

Table 1.1 Return to work rate mental health participants

Randomisation Group	Return to work rate
Health	46
Workplace	46
Combined	50
Control	59
<i>Weighted base</i>	<i>846</i>
<i>Unweighted base¹</i>	<i>645</i>

Column Percentages

¹ Base: all trial participants with Mental Health as primary health condition at time of trial entry with outcome data available.

It is not clear whether this finding is reliable or if it is a statistical anomaly. Even with this knowledge such a striking finding could not be ignored because of the disadvantage this type of intervention may have generated. The purpose of the study reported in this working paper was to test the apparent finding and to generate possible explanations for it.

1.2 The data sources

1.2.1 Quantitative data sources

Administrative data

Three administrative data sources were used. These were known collectively as the JRRP Management Information System (MIS):

- screening data collected from all those randomised during the telephone screening conducted at the beginning of the trial (854 mental health participants);
- the allocation of participants into one of the four groups recorded in the randomisation database;
- and data recorded by the Providers on their contacts with each trial participant and the interventions provided. These records were of variable quality. During the main evaluation of JRRP a randomly selected sub-sample were edited to improve the quality of the data. The records were augmented using the detailed case notes held by Providers. The editing was conducted by NatCen Only the edited sub-sample could be analysed, and contained 374 mental health participants for this analysis.

Survey data

The Outcome Survey collected data on all trial participants (including the control group) a fixed number of weeks after they first went off sick. This data was collected by interviewers face-to-face in people's homes. Participation in the survey was voluntary and not everybody agreed to take part. A response rate of 76 per cent was achieved.

A telephone survey was carried out with all members of the control group and all those screened out as ineligible (SoSOC). This survey took place before the Outcome Survey (the aim being to give an early indication of whether the screening tool was effectively screening out those at least risk of losing their jobs). The SoSOC data is used to create non-response weights to correct for non-response bias in the Outcome Survey.

Eighty-one per cent of trial entrants were interviewed, either as part of the SoSOC or as part of the Outcome Survey. The proportion of participants without follow-up data differs by randomisation group. Imputation and weighting adjustments were made at the time of the evaluation of the trial to ensure that the randomisation groups would be as comparable as possible. For further details see Purdon et al 2006.

Of the 2,845 people who entered the trial, 854 had a mental health condition as their primary health condition. Of the 2,161 participants for whom outcome data are available, 645 had a mental health condition as their primary health condition.

NOTE: The quantitative research data were not designed with separate analysis of mental health in mind. The sample sizes are often smaller than would be desirable and in some cases this prevented analysis. Furthermore, the data do not always contain the indicators that would be necessary to provide a thorough testing or explanation of the tested hypotheses.

1.2.2 Qualitative data sources

The panel study and control group study

Alongside the randomised control trial, a substantial programme of qualitative research was carried out. The main elements on which this paper draws were a longitudinal panel study of clients and a one-off study of people in the control group. (The findings of both are reported in full in Farrell et al, 2006).

The panel study included 36 participants, selected in three cohorts during the period of operation of JRRP to capture differences in experiences as services developed. The first interview was conducted face-to-face, as soon as possible after the individual had had their assessment and been assigned to an intervention group. All interviews were tape recorded, with permission, and transcribed for analysis. Five follow-up interviews were conducted by telephone at roughly monthly intervals, so that contact with participants spanned roughly six months. These interviews were tape-recorded but not transcribed: data were extracted directly into the analysis framework from the recordings.

The sample was purposively selected to ensure coverage of the following variables:

- Intervention group – 12 respondents from each of the three intervention groups
- Provider – nine respondents from four of the six locations where Providers operate
- Participant characteristics – the sampling strategy was also designed to reflect diversity in sex, age, occupation, employer type, industry sector, length of time off sick when people first contacted the service, type of employment contract (full or part time; permanent or fixed term), and main health condition

Clients continued to be interviewed regardless of the outcome of their involvement with JRRP. Twenty-seven took part in all six interviews, three missed one interview but completed the final interview, and six withdrew early from the study. The study overall consisted of 197 interviews.

The control group study comprised in-depth interviews with a sample of 46 people in the control group, to explore their experiences of sickness absence and their pathways in returning to work. The sample was purposively designed to ensure coverage of: employment outcome, age, sex, type of employer (size and sector), occupational group, main health condition and household circumstances. Interviews took place after the Outcome Survey interview. Interviews were tape-recorded and transcribed verbatim.

The data were analysed using the Framework method of analysis (Ritchie et al, 2003). This involved an analytical matrix within which there was a column for each relevant topic and a row for each interview. The data were then summarised from the transcript or recording and noted in each relevant cell.

Of the 36 panel study participants, 11 had identified mental health as their primary health condition at the screening survey. Of the 46 control group study participants, 17 had identified mental health as their primary health condition at the screening survey. The qualitative analysis reported in this working paper focuses on these two groups, comparing them with the 25 people in the intervention groups who identified a non-mental health condition as their primary health condition.

Note

The qualitative data analysis reported in this paper should be treated with caution. The qualitative research studies were not designed with separate analysis of mental health in mind. The sample sizes are smaller than would be desirable, and were not selected to provide diversity within the mental health groups. In addition, in some areas where the survey does not provide data, we report indications of prevalence in the qualitative data. Qualitative samples are not designed to support statistical analysis and these indications of prevalence cannot therefore be treated as robust – they are reported here to raise hypotheses and issues which would merit further investigation. Finally, it should also be noted that the research methods for the two studies were quite different. The panel study involved six monthly interviews beginning approximately 10 to 30 weeks after the start of sickness absence and concluding approximately 34 to 54 weeks after. The control study involved a single interview which took place up to around 75 weeks after randomisation. The panel study was designed to describe iterative developments as people used the service while the control group interviews involved looking back at a period of time which may have encouraged post-hoc rationalisation. This means that some apparent differences may be, wholly or in part, the result of the research methods used.

Other qualitative research studies

In addition to these two elements, a series of 'focused studies' was carried out looking at specific research questions of relevance to the pilot. These were:

- A study of experiences of JRRP Provider staff in establishing and delivering services, carried out through in-depth interviews and focus groups with relevant staff (Stratford et al, 2005a).
- A study of JRRP Provider organisations' perspectives on effective practices, carried out through in-depth interviews with managers and workshops with frontline staff (reported in Farrell et al, 2006).
- Two studies among employers. The first (Nice and Thornton, 2004) looked at employers' attitudes and policies towards sickness absence and used a case study design among selected employers. The second (unpublished) focused on employers who had at least one employee who had used the JRRP service and who had had some contact with the service themselves, and explored their experiences of the JRRP service.
- A study among GPs exploring how they work with patients on sickness absence (Mowlam and Lewis, 2005).

1.3 The approach taken in the reanalysis study

The approach taken in the reanalysis reported in this paper involved both descriptive and explanatory analysis. For the descriptive analysis the data were used to build up a picture of the participants to provide background information about their characteristics and circumstances, as well as exploring what happened to them in the pilot if they were allocated to an intervention.

Note

We use the following abbreviations in this paper to describe the key analytical groups:

MH-IG: people in the intervention groups who identified mental health as their primary condition at the screening stage

Non-MH-IG: people in the intervention groups who did not identify mental health as their primary condition at the screening stage

MH-CG: people in the control group who identified mental health as their primary condition at the screening stage

Non-MH-CG: people in the control group who did not identify mental health as their primary condition at the screening stage

For the explanatory analysis, a series of nine hypotheses were developed which might explain the finding that people with primary mental health conditions fared less well in the intervention group than in the control group. Most of the hypotheses were developed before reanalysis began, but some were generated by the reanalysis itself, and added and explored. This approach was taken to avoid data dredging, i.e. searching the data for significant relationships without any prior hypotheses, but also because of the small number of participants with mental health conditions and to focus on realistic and useful issues.

The nine hypotheses focused on explaining the differences in outcomes between the MH-IG and the MH-CG, and the differences between the MH-CG and non-MH-CG. They were:

H1: Despite randomisation, there are significant differences in the characteristics of the MH-IG and MH-CG: in other words, despite randomisation, it is differences between the groups which explain the statistical results rather than a real difference in job outcomes

H2: The finding reflects the definition of mental health as primary condition and is not replicated with other definitions of mental health: in other words, if we compared groups based on other definitions of mental health (such as mental health as a primary or secondary condition), the difference would not be sustained

H3: The finding reflects the definition of the primary outcome and is not replicated with other definitions: that is, it reflects the definition of a 'successful' outcome as 13 consecutive weeks in full-time work (at least 16 hours per week), including paid holiday, that occurred at or before 42 weeks after they went off sick.¹

H4: The JRRP services were more deficient for the MH-IG than for the non-MH-IG: i.e. people with mental health conditions received a less effective service than those with other conditions

H5: The quality of NHS treatment for mental conditions leaves less scope for JRRP to boost treatment for MH-IG than non-MH-IG: i.e. there was less scope for JRRP to improve outcomes for mental health clients than for those with other conditions

H6: The workplace context leaves less scope for JRRP to boost treatment for MH-IG than non-MH-IG: so that, again, there was less scope for JRRP to improve outcomes than there was for people with other conditions

H7: The pilots focused on helping clients to return to the same employer and this was differentially inappropriate for the MH-IG: so that people in this group were disadvantaged compared with the MH-CG and non-MH-IG

¹ This definition applies for participants who entered the trial after being off sick for between six and 22 weeks. For those who were off sick between 23 and 26 weeks at the time they entered the trial the return to work had to have occurred at or before 46 weeks after they went off sick.

H8: The MH-IG experienced less encouragement to return to work before they were fully fit than the MH-CG or non-MH-IG: so that returns were delayed relative to the MH-CG

H9: JRRP differentially encouraged dependency or discouraged proactivity among the MH-IG: i.e. the MH-CG took action for themselves while the MH-IG relied on the service in a way that disadvantaged them.

The first three hypotheses examine the robustness of the negative impact found for those with mental health conditions. The remaining hypotheses focus on the interventions and the environment in which they were given. It should be noted that some of the hypotheses are more powerful than others. For example, a hypothesis might be able to explain why the interventions were not able to have a positive impact on the mental health group without addressing why the MH-IG fared worse than the MH-CG.

This paper presents the descriptive analysis (Chapter 2), presents the analysis which examined each hypothesis in turn (Chapter 3), then draws together the findings in the conclusion (Chapter 4).

2 Descriptive analysis

2.1 Who are they?

The pilot was restricted to six areas of the UK and participants volunteered rather than being mandated, hence they may not reflect the national population. The participant profile given in Stratford et al 2005b illustrated some differences for the whole participant sample and it is reasonable to assume there may be differences for those with mental health conditions. Participant characteristics and circumstances were measured before they entered the pilot at the screening stage.

Health conditions were classified using ICD10.² Eight hundred and fifty four participants cited a mental health condition as the main problem that stopped them working (31 per cent). While mental health conditions vary considerably almost everyone citing a mental health condition (95 per cent) cited one of three particular types:

- reaction to severe stress and adjustment disorders (41 per cent);
- depressive episode (36 per cent);
- other anxiety disorders (18 per cent).

The severity of the condition could not be identified from the classification.

Secondary health problems were also recorded. Mental health conditions as 'other health problems' were cited by 39 per cent of participants and were not just those whose main health problem was a mental health condition.

Since the characteristics of the trial participants have been documented elsewhere (Stratford et al 2005b), this description focuses on the differences between

² ICD10 is the current version of the International Classification of Diseases code frame which is an international standard diagnostic classification used by the World Health Organisation. It allows coding of diseases and signs, symptoms, abnormalities, complaints, and injuries.

those with and without a primary mental health condition. The mental health participants were more likely to be female, 30-39 years old (and less likely to be 50-69 years old), single and working for a public company. They had higher educational qualifications and higher incomes, but also other spells off sick before the qualifying spell and non-health barriers to returning to work (particularly loss of confidence and bullying at work). There were some differences in the occupations of MH and non-MH participants. Large differences can be seen in the prevalence of recent panic attacks and whether participants had ever seen a doctor for anxiety/depression (Table 2.1), both being higher for MH participants.

Participants with mental health conditions were not recruited in the same proportion in all pilot areas. Compared to the trial as a whole, relatively more of the participants in Sheffield had a primary mental health condition, and relatively few in Birmingham and West Kent.³

Table 2.1 Comparison of characteristics for MH and non-MH participants

	<i>Column percentages</i>	
	MH sub-group	Non-MH sub-group
Recent panic attacks	69	21
Ever seen doc for anxious/depression	92	40
<i>Unweighted base</i> ¹	854	1943

¹ Base: all trial participants excluding those with missing data.

2.2 What happened in the pilot?

The Provider MIS data describe what happened to each trial participant and the interventions received (MIS data, edited sub-sample of 374 MH-IG participants). There are no equivalent records for the control group. Sample sizes were too small to examine experiences by the type of health condition for example. Nevertheless there was evidence of tailoring to participants different needs.

Some of the MH-IG did not receive a significant intervention (23 per cent), either because they withdrew after randomisation (did not consent to receive an intervention) or they were not offered anything beyond an initial assessment.⁴ This was a similar proportion to the non-MH-IG which suggests that this is unlikely to have caused the negative impact of the interventions on mental health participants.

The interventions comprised advice and treatments from a variety of people.

³ No analysis has been conducted at the area level.

⁴ Often called 'no shows'.

Restricting attention to the participants who had a chance of receiving treatments, i.e. 85 per cent of the MH-IG who gave consent to be treated, most of the MH-IG in the Health and Combined groups received treatments or treatments combined with advice, over and above the initial assessment (88 per cent and 94 per cent respectively). Fewer of the MH-IG in the Workplace group received treatments, or treatments combined with advice (66 per cent). Fifteen per cent in the Workplace group received an initial assessment only and 14 per cent received an initial assessment and advice only. This shows that far fewer participants in the Workplace group received treatment.

The length of time to complete the intervention varied considerably from under three months to more than a year. The variation was not surprising since the barriers faced by participants varied in severity and complexity. Treatments for some participants continued beyond the point at which they met the 'time off work' eligibility criteria for Incapacity Benefit despite the intention to reduce transitions to Incapacity Benefit.

The two types of treatment and advice offered in the interventions - health-related and work-related - are considered in more detail below.

2.2.1 Health-related treatments

Table 2.2 shows which health treatments were received by the MH-IG within each intervention group. The Health and Combined groups received similar levels of each type of treatment, with some exceptions. Workplace participants should not have received any health treatments; on the whole this was the case although some did receive health treatments.

Table 2.2 Health treatments received

	<i>Column percentages</i>		
	Health	Workplace	Combined
CBT	31	0	54
Counselling	44	13	38
Complementary/ Alternative therapies	37	0	21
Psychotherapy	24	0	14
Physiotherapy	12	0	10
Therapeutic exercise	10	0	10
Referral ¹	12	2	12
Other	27	8	38
<i>Weighted base</i>	<i>179</i>	<i>173</i>	<i>184</i>
<i>Unweighted base²</i>	<i>110</i>	<i>105</i>	<i>89</i>

¹ Referral to a consultant, specialist or surgeon.

² Base: all IG participants with edited MIS data.

The four most common health treatments were: Cognitive Behavioural Therapy (CBT), counselling, complementary or alternative therapies, and psychotherapy. CBT was received by more of the MH-IG in the Combined group compared to the Health group. This was partly due to a heavy focus on CBT by some of the Combined group Providers. Fewer of the MH-IG in the Combined group received psychotherapy compared to the Health group. During the pilot complementary and alternative therapies were identified by Providers as unexpectedly beneficial, and offered more than originally anticipated.

About one in ten of the MH-IG received physiotherapy which was unexpected for this sub-group. We speculate that this may have been directed at secondary health conditions rather than the primary mental health condition.

A significant proportion of the MH-IG received treatments that weren't sufficiently common to be categorised separately (labelled 'other' in the table). This category included seeing a chiropractor, Neuropsychological testing and relapse prevention planning.

2.2.2 Work-related treatments

Table 2.3 shows which work treatments were received by the MH-IG within each intervention group. The Workplace and Combined groups received similar levels of each type of treatment, with the exception of ergonomic assessments. The Workplace group received considerably more ergonomic assessments than the Combined group (37 per cent compared to 6 per cent). Our knowledge of the pilot areas suggests that the high level of provision may have been part of a treatment 'package' offered by some Providers. Health participants should not have received any work treatments, and almost none did.

Table 2.3 Work treatments received

	<i>Column percentages</i>		
	Health	Workplace	Combined
Ergonomic assessment	0	37	6
Employer liaison	0	23	24
Graduated return to work	1	7	8
Equipment bought	0	1	1
Other	1	12	16
<i>Weighted base</i>	<i>179</i>	<i>173</i>	<i>184</i>
<i>Unweighted base¹</i>	<i>110</i>	<i>105</i>	<i>89</i>

¹ Base: all IG participants with edited MIS data.

The category 'other' includes job search activities and the use of taxis for transportation to and from work.

2.2.3 Advice

It was difficult to measure and categorise the advice received from the interventions. There were three clearly identifiable themes from the advice recorded by Providers: how to return to work, how to improve health, and advice about government benefits/ debt/ financial matters. Table 2.4 shows the extent that advice was given within each intervention group.

Table 2.4 Advice received

	<i>Column percentages</i>		
	Health	Workplace	Combined
How to return to work	0	16	18
How to improve health	9	2	9
Benefits/debt/finance	6	8	17
Other	19	44	44
<i>Weighted base</i>	<i>179</i>	<i>173</i>	<i>184</i>
<i>Unweighted base¹</i>	<i>110</i>	<i>105</i>	<i>89</i>

¹ Base: all IG participants with edited MIS data.

Very little advice was received by the Health group on the three main themes. A very large proportion of the Workplace and Combined group received 'other' advice, nearly half of the MH-IG in these groups. It is not clear how this should be interpreted.

2.2.4 Why specific interventions were offered

The Providers recorded their justification for each treatment given to each participant. More than one reason could be offered for each treatment; the three most common are shown in Table 2.5. For treatments provided to the Health group equal emphasis was placed on these three reasons. For the Workplace group the most frequently cited reason was 'not available quickly', and for the Combined group 'not available locally' as well as 'not available quickly'.

Table 2.5 Main reasons for offering treatments

	<i>Column percentages</i>		
	Health	Workplace	Combined
Not available quickly	56	56	75
Not available locally	50	30	68
Person would have to pay	54	30	48
No reason given	13	48	5
<i>Weighted base</i>	<i>179</i>	<i>173</i>	<i>184</i>
<i>Unweighted base¹</i>	<i>110</i>	<i>105</i>	<i>89</i>

¹ Base: all IG participants with edited MIS data.

The table also shows that no justification was given for nearly half of the treatments received by the MH-IG in the Workplace group. It is not clear whether this is meaningful or is a result of poor record-keeping.

2.2.5 Contact as an intervention

The amount of contact with the Provider varied across participants from an average of once per month to more than ten times per month. The MH-IG in the Workplace group received less contact on average.

There were many reasons for contact, the main ones being (a) to offer support or review progress, (b) to give advice, and (c) while receiving treatment. Discarding (b) and (c) because alternative measures of advice and treatment have already been considered, of most interest were the contacts made for support/ progress review. This was the best measure of the case manager's contribution to the intervention. For two thirds of the MH-IG in the Health or Combined groups support/ progress review made up no more than 25 per cent of contacts. For the MH-IG in the Workplace group a larger proportion of contacts were for support/ progress review (almost half had no more than 25 per cent of this type of contact, and almost half had 25-50 per cent of contacts of this type).

2.2.6 Summary

Not all MH participants randomised to an intervention consented to receive it, and some consenting participants did not receive any treatments. The extent of this behaviour did not appear to jeopardise the trial. Some common health and work-related treatments could be identified although there was also evidence of tailoring to the needs of the participants. Treatments had to be justified and Providers most frequently cited reasons were that treatments were not available quickly or locally, or that the participant would have to pay to receive it otherwise. Overall there were no systematic deficiencies identified (from a non-clinical perspective).

3 Explanatory analysis: the hypotheses

3.1 H1: despite randomisation, there are significant differences in the characteristics of the MH-IG and MH-CG

The design of the pilot evaluation allowed a direct comparison of the outcomes of each intervention because each group was expected to contain participants with the same characteristics, on average. However since the randomisation process did not control for health condition there could be differences in the characteristics of the MH-IG and MH-CG, purely by chance. Any differences have the potential to affect the impact estimate.

Demographic, health-related and work-related characteristics were measured before participants entered the pilot at the screening stage. The MH-IG and MH-CG were closely matched on these observable characteristics. The only notable difference found was higher qualifications on average for the MH-CG. This suggests that there isn't enough evidence to support hypothesis H1; we conclude that the randomisation was robust enough.

However further differences were found for the weighted sub-group of survey respondents. Since the outcome measure was based on weighted survey respondents this suggests that there is merit in further exploration of the outcome measure (discussed in relation to hypothesis H3 in section 3.3).

3.2 H2: the finding reflects the definition of mental health as primary condition and is not replicated with other definitions of mental health

Mental health participants were defined as those declaring at the screening stage that the main problem that stopped them working was a mental health condition. The definition determines who to include in the analysis, so that changes could affect the impact estimate. There were two concerns about the definition: first that conditions other than the main (or primary) condition may have been important with respect to returning to work, and second that conditions may not have been measured accurately. The potential for alternative definitions was explored. Where feasible alternatives could be derived, impact estimates were calculated for comparison.

Two new definitions were developed to address the first concern described above. The definitions expanded the original definition, adding either (a) participants with secondary mental health conditions or (b) participants recently experiencing panic attacks, (all recorded at the screening stage). The latter definition was not an indicator of a diagnosis but had the potential to act as a proxy for having a condition that affected return to work. The return to work rates calculated for the MH-IG from the two new definitions were of a similar magnitude to the original and the MH-CG had higher returns to work than the MH-IG i.e. same trend exhibited by the original definition.⁵ The difference between the MH-IG and MH-CG for the definition that included participants recently experiencing panic attacks was relatively small. While this lends some support to the hypothesis being tested the credibility of including panic attacks is questionable.

A third new definition was derived to explore the second concern (that conditions may not have been measured accurately), by including participants only if their health condition was validated by the Provider (a reduction in the number of mental health cases identified compared to the original). In the original definition the conditions were described by the participants and recorded by interviewers for subsequent coding to ICD10. Inaccuracies could have arisen because ICD10 is a clinical tool but the information used was not a professional clinical diagnosis, interviewers may have recorded conditions erroneously because they were not trained medical experts, and the participants may have been reticent disclosing their conditions. Providers had medical expertise, were more likely to be able to accurately assess what condition was the main problem relating to being off work and more contact with participants. Thus their recording was hypothesized as being more accurate. This new definition could be calculated for the MH-IG only (Providers did not have any contact with the MH-CG). Again the impact estimate calculated was of a similar magnitude to the original.⁶

⁵ Adding secondary conditions: MH-IG = 46%, MH-CG = 55%. Adding panic attacks: MH-IG = 44%, MH-CG = 48%.

⁶ Provider validated: MH-IG = 45%.

One final piece of evidence here is the fact that there was generally a very good fit between the qualitative data and the quantitative screening data in the identification of mental health as the primary condition. Only one of the 28 people in the panel study or control group study identified by the screen as having mental health as a primary condition did not describe their health clearly in these terms. This was a woman with high blood pressure which she attributed in part to stress and worry, but who discussed her health in terms of a physical rather than mental health condition.

Thus we conclude that there is no evidence to support hypothesis H2 and the original definition of mental health can be retained.

3.3 H3: the finding reflects the definition of the primary outcome and is not replicated with other definitions

The impact estimate was based on the primary outcome measure of return to work. A participant was considered to have been successful if they had returned to full-time work by week 29/33⁷ after going off sick, without any break in employment since going off sick, and that it continued for at least 13 weeks. This measure of success was derived at the outset of the pilot reflecting policy considerations and service provision incentives. The definition was broken down into its constituent parts to explore whether any element might be responsible for the negative impact. Other related elements were also explored: survey non-response and other outcomes. Each element is considered in turn below.

3.3.1 Full-time

Full-time work was measured as 16 hours or more per week so we tested whether a definition of fewer hours would alter the observed difference between the MH-IG and MH-CG. Redefining a success to include working fewer hours resulted in an impact estimate that shows the same trend as the original definition.⁸

3.3.2 By week 29/33

The reference period for the return to work may have been too short for the impact of the interventions to have been visible. Without further data collection efforts (a follow-up survey or matching to other administrative sources) the outcome cannot be replicated for a longer reference period. Rather than ignoring this element of the definition, the start dates of returns to work were examined to give an indication of any differences between the MH-IG and MH-CG.

⁷ If the participant had been off sick for 6-22 weeks at the time they entered the pilot. Extended to week 33 if the participant had been off sick for 23-26 weeks.

⁸ Allowing any hours: Health = 46%, Workplace = 47%, Combined = 51%, Control group = 57%.

Work start dates show that over the weeks up to week 29/33 the MH-CG return to work earlier than the MH-IG. The largest difference was found at week 14 where around 15% more of the MH-CG had started a return to work than the MH-IG. The difference decreased in later weeks. If the MH-IG were delayed by the interventions then a longer reference period might have shown an increased return to work rate. Therefore the evidence does not allow us to discount the possibility that the reference period might have affected the impact estimate.

3.3.3 No break in employment

The requirement that a contract of employment was to be maintained throughout without a break may have been challenging for some participants, particularly when people did not return to the same employer it may have been hard to secure an alternative without a break. Since employment contracts were not recorded explicitly in the survey work histories the outcome cannot be redefined to examine this element. A related discussion can be found in section 3.7 however.

3.3.4 13 continuous weeks

The requirement to work continuously may have been too difficult for participants at the outset of their return to work. If returns were more erratic they would have been missed from the measure. This was examined by calculating the impact for returns of shorter continuous durations. The same trend of a higher return to work rate for the MH-CG compared to the MH-IG was seen for all durations (from one to thirteen weeks). Testing showed significant differences for the longer durations only (10+ weeks). Therefore this element of the outcome definition is unlikely to have affected the impact estimate.

3.3.5 Survey non-response

The outcome was measured using survey data which was subject to non-response. During the pilot evaluation this was examined and a weighting strategy adopted to try to correct for any bias introduced by differential non-response. If the non-response weighting correction was not appropriate for mental health participants, then the impact estimate might be biased.

The effect of the weighting was explored by re-calculating the impact estimate using only one of the two original weighting components.⁹ The omitted component had made adjustments to the MH-CG but did not affect the MH-IG. This was an obvious candidate for exploration since the return to work rate of the MH-CG was so much higher than the rate for the non-MH-CG. The same trend is seen for the

⁹ The full weight had two components (see Purdon et al 2006 for full details). The first component made adjustments to IG and CG for participants without any outcome data, the second component adjusted the CG only for participants who had responded to SoSOC but not the Outcome Survey (had work history information relating to their earliest weeks in the pilot but not full outcome data). The investigation relates to the second component.

'partially' weighted data, although the difference has been reduced (Table 3.1, first two columns). The omitted weighting component appears to magnify MH-CG participants who returned to work and because this component does not get applied to the MH-IG it is an uneven adjustment.

Table 3.1 Return to work rate for mental health participants: different weights

Randomisation group	Column percentages		
	Return to work rate	Return to work rate	Return to work rate
	Full weight	Partial weight	Optimal weight
Health	46	46	46
Workplace	46	46	46
Combined	50	50	49
Control	59	55	58
<i>Weighted base</i>	<i>846</i>	<i>848</i>	<i>854</i>
<i>Unweighted base¹</i>	<i>645</i>	<i>645</i>	<i>645</i>

¹ Base: all trial participants with Mental Health as primary health condition at time of trial entry with outcome data available.

In order to understand how omitting the weighting component could reduce the difference in return to work rates we return to the original survey data. The MH participants had different non-response behaviour to the non-MH participants which was not accounted for in the weighting. The weighting assumed that non-respondents to the Outcome Survey had higher return to work rates than respondents (as estimated from SoSOC). This was true overall and for non-MH participants, but was not true for MH participants. The return to work rates for MH participants were similar for respondents and non-respondents to the Outcome Survey. Therefore part of the weighting inappropriately magnified more successful MH participants, and removing this component reduced the return to work rate of the MH-CG. An optimal weight was derived, replacing the omitted component with an appropriate revised version (Table 3.1, final column). The optimal weight produces results that are very similar to the original results.

This issue was explored further by testing the differences in the characteristics of the MH-IG and MH-CG for weighted survey respondents. Similar characteristics and testing was used to those used to test hypothesis H1; further significant differences were found. More complex analysis techniques were applied to re-estimate the impact while trying to account for the differences in characteristics (propensity score matching and regression analysis).¹⁰ Although these techniques are usually reserved for larger samples, the samples were sufficient to obtain an

¹⁰ Conducted by John D'Souza.

indication of the effect of the characteristics. They showed that any differences in characteristics between the participants in each of the randomly allocated groups do not appear to be responsible for the negative impact of the pilot interventions. The large difference in return to work rates remains.¹¹

3.3.6 Other outcomes

The interventions were directed towards helping participants return to work. However they may have affected other aspects of participants lives such as health or future work expectations. Comparing the MH-IG and MH-CG, there were no significant differences in outcomes relating to receiving Incapacity Benefit, doing voluntary work, education, training, or their financial training, as reported in the survey. There were significant differences in health outcomes: the Health and Combined interventions appeared to have significantly improved participants self-reported health compared to the Workplace and Control group participants.¹² Also a greater proportion of the non-MH participants still had their health condition (explored further in hypothesis H8).

The data collected about future work expectations indicated that fewer non-MH participants who were not working were looking for work, than MH participants. This might help to explain the difference in the return to work rates of the MH-CG and non-MH-CG. Perhaps it was not that the MH-CG found it easier to return to work but that more of them were trying than the non-MH-CG.

One possible explanation would have been if the aspirations of the MH participants were not directed towards a return to work. The qualitative data showed that in all sub-groups there were some other objectives than returning to work, particularly wanting to improve health or avoid job termination but there was nothing to suggest that return to work was an inappropriate focus.

3.3.7 Conclusion

The reference period used to measure the outcome may have had some effect but the definition of the primary outcome is unlikely to explain the large negative impact. The differences in characteristics identified from hypothesis H1 although magnifying the impact do not seem to explain the negative impact adequately either. Differences in work aspirations may be responsible for the success seen for the MH-CG but further data would be needed to explore this fully.

¹¹ Return to work rates from propensity score matching: Health = 36%, Workplace = 50%, Combined = 48%, Control group = 59%.

¹² Health measures showing significant differences: self-assessment of general health and HAD anxiety scale.

3.4 H4: the JRRP services were more deficient for the MH-IG than for the non-MH-IG

The Provider's records of interventions allow an exploration of the quality, in a general sense, of the service received by the intervention groups. Similar proportions of MH-IG and non-MH-IG received treatments of some sort. Note that small sample sizes have restricted the comparisons, and that this is not an assessment of the adequacy of treatments in a clinical sense.

Looking at the details of the treatments received there are clear differences between the MH-IG and the non-MH-IG. For example, the MH-IG received less physiotherapy, fewer referrals to consultants and ergonomic assessments but more counselling and CBT. This was to be expected. There were no consistent patterns across all intervention groups.

A number of general measures were created from the quantitative data to compare the MH-IG against the non-MH-IG in terms of range, intensity and reason for treatment as well as the satisfaction ratings given by participants. The measures were not consistently different. For the Health intervention compared to the non-MH participants more of the MH participants received nothing (and perceived this) although when treatments were received they were relatively more intense. For the Combined intervention compared to the non-MH more of the MH participants received treatments, and these were from a broader range and more intense. There were no significant differences for the Workplace intervention.

Similarly, there was no evidence from the qualitative panel study to suggest that the JRRP services were more deficient for the MH-IG than for the non-MH-IG. Both groups contained both people who were highly satisfied with the help they received and people who were dissatisfied. Where people in the MH-IG were not satisfied with the help they received this related to either a form of help they wanted not being available (for example one person wanted workplace support but was assigned to the health group; one wanted help with stress management which was never offered; one wanted a minor operation for a physical condition which the pilot did not fund), or to finding the assessment or psychological therapies uncomfortable or intrusive. The same issues, and other sources of dissatisfaction, were noted by people in the non-MH-IG.

From a non-clinical perspective, we conclude that there is no evidence to support hypothesis H4.

3.5 H5: the quality of NHS treatment for mental conditions leaves less scope for JRRP to boost treatment for MH-IG than non-MH-IG

There is evidence about the services received outside JRRP from the qualitative data and to a much lesser extent from the survey data sets.

All the people in the qualitative research MH-IG sample had been offered some form of medical intervention by their GP, either medication, counselling, referral to psychologists or psychiatrists or intensive support from a Community Mental Health Team. This does not mean it was necessarily successful. Some people turned down medication or other support, either because they did not want it or because they were already accessing counselling from the JRRP Provider or from another source. Treatment for counselling was sometimes delayed by long waiting lists. People also sometimes found that anti-depressant medication either made no difference or had unpleasant side effects. However, everyone had had a diagnosis and been offered some form of help: none had a condition that had not been identified or for which treatment was not available. The situation was a little more mixed for the MH-CG, where there were some who had not been offered either medication or any form of talking therapy by their GP.

However, the picture was very different in the non-MH-IG qualitative sample. Here, there was a clearer pattern of people having no NHS treatment or treatment for some symptoms but not for their main condition. This arose in two ways. The first situation was where people had not yet had a diagnosis so that treatment could not be given. They were either on a waiting list for investigative intervention, or initial consultations had not been successful and they were awaiting further investigative interventions, or they had reached the end of the diagnostic interventions available and had not had a definitive diagnosis. The second situation was where people had been told that there was no treatment for their condition. People thus described a frustrating situation where they had become 'stuck' within the NHS, either on a waiting list or having exhausted their options. Musculo-skeletal conditions were common here (arthritis, spondylitis, sciatica, other back problems, conditions affecting hands) but it also arose where people had chronic fatigue syndrome, epilepsy, strokes or bowel conditions.

The survey data show very few participants recalled accessing health advice or treatments organisations other than JRRP. In the four weeks before the survey interview around half of the MH and non-MH participants had talked to their GP. Within the same period more MH participants had talking therapy than non-MH participants (20 per cent compared to three per cent) and this was more prevalent for those in the Workplace and Control groups.

For people with mental health conditions in the qualitative sample, JRRP was able to provide treatment additional to their NHS treatment by providing faster access to psychological therapies, additional psychological therapy to what people were already receiving, addressing other (physical) conditions or providing exercise support or complementary therapy.

However, there seemed to be more scope to add to NHS treatment for people with non-mental health conditions. The form this took varied and involved:

- faster access to investigative interventions;
- faster access to minor surgical interventions;

- physiotherapy (either which was not being offered or which was compared favourably with NHS provision in terms of quality or intensity), pilates or support with exercise;
- complementary therapies;
- condition management or occupational health advice.

There were still people who the service was not able to help, particularly if they were awaiting expensive treatment. But there seemed to be more scope for the service to boost NHS provision than for the MH-IG¹³.

We conclude that there is some evidence to support hypothesis H5.

3.6 H6: the workplace context leaves less scope for JRRP to boost treatment for MH-IG than non-MH-IG

The quantitative data cannot give a full picture of the workplace context before participants started the pilot. However some indicators of the supportiveness of the workplace were collected. At the screening stage the MH participants were more likely to say that they didn't like their job (first two rows in Table 3.2), and that their employer was not very understanding about them being off sick. Additionally the MH participants were more likely to have experienced bullying at work.

Table 3.2 Measures of supportiveness of the workplace

	<i>Column percentages</i>	
	MH	Non-MH
Doesn't like job	19	6
Doesn't know if likes job	14	6
Employer not understanding	32	12
Bullying at work	12	1
<i>Unweighted base¹</i>	<i>854</i>	<i>1946</i>

¹ Base: all trial participants with screening data.

The qualitative data shows a mixed picture in whether or not work was implicated in mental health conditions. Within the MH-IG and MH-CG, there were people who saw their condition as having been caused by work, for example arising from

¹³ In the qualitative study among GPs (Mowlam and Lewis, 2005), GPs noted the absence of a diagnosis as a barrier to patients' recovery and the long waiting lists for diagnostic interventions as well as for counselling. However the study does not provide direct evidence either to support or to refute the hypothesis that the quality of NHS treatment for mental health conditions leaves less scope for JRRP to boost treatment.

bullying, unfair treatment, pressure, being poorly supported, reorganisation that left people feeling undermined or isolated, or their physical working conditions.

However, across the MH-IG and MH-CG it was rare for jobs to have been terminated or for termination to have been threatened by the employer. Apart from this, how supportive employers appeared to be varied. There was a group of cases where there was little or no contact with the employer, a group where a line manager was not supportive but someone else in the organization was (particularly in Occupational Health), or where redeployment was being discussed, and a group where the participant described a very supportive and non-pressurising approach being taken by the employer.

Among the non-MH-IG the picture was rather different and there was a clearer pattern of negative contact with the employer. There was higher prevalence of cases where the job had been terminated by the employer or where the employer was threatening the post, for example setting a deadline for a return to work that the person was unlikely to meet or discussing termination. There were also cases where the individual was trying to negotiate a return but was experiencing either non-response from the employer or an inflexible approach, being told that light duties, a phased return or redeployment were not possible.

The hypothesis from the qualitative data is thus that, although the survey shows a higher proportion of people with mental health conditions did not like their job, employers were somewhat more supportive of a return to work than where there were physical conditions¹⁴. The quantitative data offers weak support for these differences, although we do not have an overall measure of perceived employer attitude or support. Compared to the non-MH group more of the MH participants (particularly the MH-CG) accessed help from their workplace Occupational Health Department (42 per cent compared to 34 per cent) and help from their Union (27 per cent compared to 19 per cent).

To summarise, there is mixed evidence for hypothesis H6; the MH participants seem to have more problems with the job they were doing but there appeared to be more scope for other support or solutions.

3.7 H7: the focus on returning to the same employer was inappropriate for the MH-IG

Looking back at how the pilot was implemented it seems that participants were expected to return to the same employer. This was not stated explicitly but no breaks in employment were allowed. That is not to say that different employers were not considered - job search was one of the services offered as a work-related treatment.

¹⁴ The studies with employers (including Nice and Thornton, 2004) do not provide evidence which either supports or refutes the hypothesis that employers were less likely to terminate or threaten contracts for employees with mental health conditions

The survey data show whether participants' returned to work and, if so, if this was with the same or a different employer. Similar proportions of the MH and non-MH participants were not working or were working for the same employer, but a significantly higher proportion of the MH participants were working for a different employer (15 per cent compared to 7 per cent). This was not surprising given the higher levels of dislike for their job (see section 3.6). The move to a different employer was more common for the MH-CG than the MH-IG (22 per cent compared to 13 per cent). We turn to the qualitative data to uncover possible explanations about how a return to work with a different employer came about.

From the studies among Provider staff, it was clear that the focus of the services was on a return to the same employer, and the discussion about job outcomes in the staff research workshops and interviews was almost entirely about returns to the same employer rather than about new employers. This emphasis was explained by both the perceived status of a return to the same employer as a 'successful' outcome and the view that it was easier for a client to return to the old job because there is an existing relationship. The quality of support to look for other employment was generally criticized, by participants but also to some extent by Provider staff.

Participants too generally saw advantages to returning to the same employer. Where an employer had been supportive or at least kept the job open they said they would feel guilty about finding another job instead. There was a sense of 'better the devil you know' than the risks involved in a new job, and people also talked about not wanting to jeopardize pension rights. People felt unconfident about finding a new employer, particularly given their health record but also if they were older or otherwise felt less able to compete in the jobs market. Finally, one person who had been bullied talked about wanting to prove to their employer that they had not been defeated, by returning to the same job. These views emerged both among people with mental health conditions and those with other health conditions.

In practice, very few people in the qualitative samples actually got a new job, although others contemplated it. The data suggest that people were more likely to get or contemplate a new job in the non-MH-IG and the MH-CG compared with the MH-IG. The focus on another employer, where it occurred, appeared to be prompted by two things. First, the employer had terminated the job or was threatening termination. Second, the participant had reached the view that their previous job was incompatible with their health. Jobs were seen as incompatible with health among people in the MH-CG in some cases where the work was seen as the cause of the condition. However, in the qualitative data there was a higher prevalence among the non-MH-IG group of people concluding that their condition was incompatible with their job and that they would not recover sufficiently to be able to return – a conclusion they reach themselves or that was prompted by their employer or GP. In essence, it seemed that people with physical conditions got a clearer reading that they would not be able to return to their job than people with

mental health conditions. But they were not necessarily able to act on that reading by looking for another job because they were not well enough yet to contemplate any return to work.

The evidence suggests that hypothesis H7 is feasible, at least for some people with mental health conditions, although the available data is limited.

3.8 H8: the MH-IG experienced less encouragement to return to work before they were fully fit

Provider staff emphasized that people did not have to be fully fit to return to work. The extent that this was possible was likely to depend on how their health condition restricted them as well as on the demands of their job, the scope for temporary re-deployment and commitments within their home environment. Exploring the timing of returns to work in relation to the recovery of full capacity is a complex issue. The potential interactions between these factors may be important and can be considered simultaneously within the qualitative data analysis. The quantitative analysis is limited, partly because of the restricted sample sizes.

The qualitative data suggest that people with mental health conditions experienced less pressure to return to work than those with other conditions. This arose in two ways. First, as we have seen, the jobs of people with mental health conditions in the qualitative research samples were less likely to be terminated or threatened by employers. Second, there seemed to be more messages encouraging caution and delay in returning to work for people with mental health conditions, and more encouragement to return to at least some work before they were fully fit among people with physical health conditions.

All but one of the MH-IG reported either family members, a GP or their employer strongly encouraging them not to return to work until they were sure they were ready. Although people in the non-MH-IG also described similar messages of caution, there were also cases where people described their GP refusing them further sick notes, encouraging them to return to work, or initiating discussions about returning to work on a phased return or light duties. There were also cases where employers were pressurizing the person to return to work or, as we have seen, threatening to dismiss them if they did not. People in the non-MH-IG also described more encouragement from Providers to return to work part-time or on light duties; there was very little evidence of this type of encouragement by Providers in the MH-IG sample. A complementary source, the MIS, shows a mixed picture where neither sub-group received significantly more of all of these types of treatment. However these treatments were not very common.

The survey data measured how well they were feeling generally as well as the improvement in the primary health condition that had been stopping participants working, at the time of the survey (42/46 weeks after going off sick). This is the best available measure because during their health status during the pilot was not recorded in a way that allows a precise measure of this hypothesis. Participants

with successful return to work outcomes gave significantly higher assessments of their health.¹⁵ This was the case for all participants regardless of whether they had a mental health condition or not and whether they were in an intervention group or the control group. For the hypothesis to be accepted the data need to show that the MH-IG have been treated differently from the non-MH-IG. A comparison of the health of the MH-IG and non-MH-IG with successful outcomes shows that fewer of the MH-IG still have their primary health condition than the non-MH-IG (62 per cent compared to 74 per cent) and they generally have better health (more 'good' and fewer 'fair' or 'poor' for MH-IG). The same trend is seen for the control group, so it looks as though participants with MH conditions in both the intervention and the control group delayed their return to work until they were fully fit, not just those in intervention groups.

The evidence suggests that hypothesis H8 is feasible.

3.9 H9: JRRP differentially encouraged dependency/ discouraged proactivity among the MH-IG

From the qualitative data, there is some evidence to support the hypothesis that people in the MH-IG were particularly dependent on the services and less likely to act for themselves than those in the MH-CG. Although these were not labels that people applied to themselves, and we should therefore be cautious in our interpretation of their behaviour, there was a clear pattern of behaviour among people in the MH-IG that could be seen as suggesting dependency:

- where people received intensive treatment, sometimes overlapping with other provision, and described this as being driven by the Provider rather than by them. For example, one person had what appeared to be at least 18 sessions of CBT, and twice had assumed the sessions were coming to an end but was encouraged to have more by the Provider;
- where people saw the Provider as being in control of their treatment plan and said that they were happy to go along with whatever the Provider offered without seeming to form a view themselves about what treatment was needed;
- where people seemed to be reluctant to form their own view or make their own decisions about whether they could return to work and instead said they would take the advice of their Provider or GP;
- where people needed the close involvement of a member of staff to progress agreed activities as part of their treatment plan, for example being encouraged to do worksheets or swimming sessions alone but not doing them unless they were accompanied by a staff member;

¹⁵ The data shows that health and return to work are linked but cannot determine the direction of the effect, that is, whether better health leads to return to work or return to work leads to better health.

- describing a particularly strong attachment to their case manager (and similarly to other people such as GPs);
- not initiating contact themselves and relying on the Provider to keep in touch, so that if the Provider did not maintain contact they lost touch with the service.

There were also examples of similar behaviours among people in the non-MH-IG, and examples of more proactive behaviour among the MH-IG such as initiating contact with Providers or employers and discussion about returning to work. However, in the qualitative data examples of passive behaviour were particularly evident among the MH-IG.

This contrasted starkly with the MH-CG. Again, there were examples of what might be characterized as both dependent or passive and more proactive behaviours. However, there was a clearer pattern of proactive behaviour among the MH-CG. For example, people initiated contact with the employer; initiated discussions about returning to work or lobbied actively for redeployment; used other sources of support such as Jobcentre Plus, employee assistance provision, trade unions, private counselling or other treatment or accessing training; or pressed their GP or other health professionals for treatment or to support a return to work.

There were also occasionally cases in the MH-CG where people said that being turned down for the service made them realise that they would have to act for themselves and could not rely on help from anywhere else. However, this was also found among the non-MH-CG and there was nothing to suggest that mental health was a factor. However one of the principles underlying the design of the pilot was that participants would behave normally. In other words, there was an assumption that any disappointment arising from being allocated to the control group would not induce them to put in any more or less effort than they would have done otherwise. If the experiment-induced attitude described here was widespread amongst the MH-CG it could have generated more effort, leading to more returning to work. Unfortunately there is no measure of the prevalence of this attitude or any resulting changes in behaviour.

There was very little useful data from the quantitative sources for the exploration of dependency or experiment-induced effects. What data there are does not support what was found in the qualitative sources. The MH-IG and non-MH-IG spent the same length of time receiving treatments on average, and showed the same extent of initiating contact with the service (13 per cent).

One final issue from the qualitative data is that the MH-CG also appeared to have clearer views about what they would have wanted from the JRRP service. People in the MH-IG sometimes had clear ideas but generally, like people in the non-MH-IG, they expressed quite vague aspirations. They talked about it as something that 'could do no harm', and said that at this stage they would have tried anything that might help or 'would have clutched at any straw', or talked more generally about wanting help to get better and to get back to work. In the control group, however, people had more specific ideas about what they would have wanted

from the service – particular forms of health treatment, help negotiating with the NHS or different forms of workplace support.

This evidence needs to be treated with caution not only because of the small samples and the fact that the study was not design to support quantitative analysis but also because of the different research methods used with the panel and the control group samples. The panel study was designed to describe iterative developments as people used the service while the control group interviews involved looking back at a period of time which may have encouraged people to foreground themselves as active players. It may also be that it was harder for people who subsequently had more contact with the services to remember their initial motivations, whereas being refused support in some way crystallised the aspirations of people in the control group or gave them a stronger feeling about what they had missed out on.

Finally, it is perhaps also worth noting that the Provider staff acknowledged the potential for clients, and particularly clients with mental health conditions, to become dependent on the service. They described aiming to shift the locus of control and action from themselves to the client, encouraging the client to take more responsibility as time went on. However, the qualitative sample of MH-IG cases suggests that people remained largely dependent on the Provider to initiate contact.

There was some evidence of dependency among the MH-IG, supporting hypothesis H9.

4 Conclusions

In this study we explored the negative impact of JRRP interventions for people with mental health conditions. A number of possible explanations, our hypotheses, were tested using quantitative and qualitative data. Several of the hypotheses were supported by the data.

Three out of the nine hypotheses tested whether the negative impact was likely to have been caused by design-related issues. Two of these were discarded outright (H1 and H2). The third (H3) indicated that the reference period used in the measurement of success and non-response to the measure, may have had some effect but is not believed to be capable of causing such a large negative impact. The corrective non-response weighting applied to the survey data (H3) magnified an effect that was already there rather than generating the effect.

Five of the remaining six hypotheses were found to offer plausible explanations:

- returns to work appear to have been delayed for the MH-IG for three reasons: focus on the same employer where it may have been more productive to seek a new job (H7), waiting for a more complete health recovery (H8), and dependency on the Provider (H9);
- there appeared to be less scope to boost NHS provision for mental health conditions compared to other conditions (H5);
- there was some evidence of employer behaviours that were more supportive (or less obviously obstructive) of a return to work which meant less scope to improve the workplace context for mental health (H6).

The exploration suggested that there could be a combination of factors responsible for the negative impact. Hypotheses H5 and H6 explain why the return to work rate for the MH-CG was higher than the non-MH-CG, and hypotheses H7, H8 and H9 explain why the return to work rate was lower for the MH-IG than the MH-CG. There was no evidence to indicate that the interventions were more deficient for the MH-IG than for the non-MH-IG (H4).

References

Farrell C, Nice K, Lewis J and Sainsbury R (2006) Experiences of the Job Retention and Rehabilitation Pilot DWP Research Report No 339 London: Corporate Document Services.

Mowlam A and Lewis J (2005) Exploring how General Practitioners work with patients on sick leave DWP Research Report No. 257 London: Corporate Document Services.

Nice K and Thornton P (2004) Job Retention and Rehabilitation Pilot: Employers' management of long-term sickness absence DWP Research Report No. 227 London: Corporate Document Services.

Purdon S, Stratford N, Taylor R, Natarajan L, Bell S and Whittenburg D (2006) Impacts of the Job Retention and Rehabilitation Pilot DWP Research Report No 342 London: Corporate Document Services.

Stratford N, Farrell C, Natarajan L and Lewis J (2005a) Taking Part in a Randomised Control Trial: a participant's eye-view of the Job Retention and Rehabilitation pilot DWP Research Report No. 273 London: Corporate Document Services.

Stratford N, Taylor R, Legard R, Natarajan L, Purdon S and Shaw A (2005b) The Job Retention and Rehabilitation Pilot: reflections on running a randomised controlled trial DWP Research Report No. 305 London: Corporate Document Services.

Ritchie J, Spencer L and O'Connor W 'Carrying out Qualitative Analysis' in Ritchie and Lewis (eds) (2003) Qualitative Research Practice Sage: London.

