Prospective pilot evaluation of the effectiveness and cost-utility of a 'health first' case management service for long-term Incapacity Benefit recipients.

J Warren[1,2], C Bambra[1,2], A Kasim[2], K Garthwaite[1,2], J Mason[2,3], and M Booth[2,3]

Affiliations:

1. Department of Geography, Durham University

2. Wolfson Research Institute for Health and Wellbeing, Durham University

3. School of Medicine, Pharmacy and Health, Durham University

Corresponding author: Professor Clare Bambra, Department of Geography, Wolfson Research Institute for Health and Wellbeing, Durham University Queens Campus, Stockton on Tees, TS17 6BH UK Tel: 0191 334 0417 Email: clare.bambra@durham.ac.uk

Keywords: primary care, welfare, worklessness, sickness absence

Word count: 2997

Abstract

Background: In line with NICE guidance, an NHS commissioned case management intervention was provided for individuals receiving Incapacity Benefit payments for three years or more in the North East of England. The intervention aimed to improve the health of participants.

Methods: 131 participants receiving the intervention were compared over nine months with a (non-equivalent) comparison group of 229 receiving Incapacity Benefit payments and usual care. Health was measured using EQ-5D, EQ-VAS, SF-8 HADS and the Nordic Musculoskeletal questionnaire. Socio-demographic and health behaviour data were also collected. Fixed effects linear models with correlated errors were used to compare health changes between groups over time. A preliminary cost-utility analysis was also conducted.

Results: Comparison group measures of health were stable over time. Starting from comparatively poor initial levels, case management group generic (EQ5D, EQ-VAS) and mental health (HADS-A, HADS-D and SF8-MCS) measures improved within six months to similar levels found in the comparison group. Musculoskeletal (Nordic 2) and health behaviours did not improve. Tentative estimates of cost-utility suggest the intervention cost in the region of £16,700 to £23,500 per QALY.

Conclusion: Case management interventions may improve the health of incapacity benefit recipients. Further research is required to help confirm these pilot findings.

200 words

Background

Poor health is a significant risk factor for worklessness, as well as remaining out of work [1]. In most advanced market democracies, long-term health-related worklessness carries an entitlement to receipt of financial support from the welfare state in the form of sickness and disability payments or, in the case of the UK, incapacity-related benefits (Incapacity Benefit or Employment and Support Allowance). Rates of receipt of these benefits have increased from 0.5 million recipients in 1975 to 2.6 million in 2007 - around 7% of the UK working age population, accounting for 11% of UK social security expenditure or 1.8% of gross domestic product (GDP) [2].

This is an increasingly prominent policy concern in the UK with various welfare-to-work interventions initiated since the 1990s (for an overview see [3]). The effectiveness of such interventions in increasing labour activity amongst incapacity-related benefit (IB) recipients has been questioned by numerous evaluations [4;5;6]. A number of different reasons have been suggested for the lack of effectiveness, including lack of demand side interventions [5], scarcity of jobs [7], and the lack of attention paid to the barriers to employment imposed by ill-health [8;9]. The latter has led to calls for welfare to work interventions to take more of a 'health first' approach [8].

Additionally, the 2008 Black review of the health of the working age population has played a key role in initiating a wider debate about work and health in general [10]. This resulted in the introduction of a "Fit Note" to general practice to replace the traditional sick note as well as the initiation in 2009 of eleven "Fit for work" pilots across the UK (seven of which are still running and will continue until March 2013) [10] which aim to test the impact of biopsychosocial case management interventions on return to work [10]. This was followed up by a comprehensive review of sickness absence arrangements to identify ways of reducing labour market drop-out as a result of ill health and sickness absence [11].

In 2009, the National Institute for Health and Clinical Excellence (NICE) produced guidelines on the management of incapacity and sickness absence [12; 2] which recommended that case management approaches were the most effective in achieving return to work (Box 1). This paper presents results of an evaluation of a the health effects (with cost-utility estimates) of a pilot case management service

for long-term IB recipients commissioned by a Primary Care Trust (PCT) in the North East of England in 2009 (where 8.4% of the working age population receives IB). The intervention was informed by the NICE guidance and although it pre-dated the Black Review [10] and the Independent Review of Sickness Absence [11], the evaluation can make a timely and valuable contribution to the current policy debate about workplace health and sickness absence.

Method

Study Design

The case management intervention and a comparison group were compared prospectively in a non-equivalent group, repeated measures design using questionnaires at base line (T1), 3 months (the intervention's midpoint, T2), 6 months (the intervention's endpoint, T3) and 9 months (three months post-intervention, T4). Within the pilot design, differences between groups were explored using a range of generic and disease-specific outcomes as well as health behaviours. Tentative retrospective cost-utility analysis was also conducted.

Case Management Intervention Group

The service was delivered by a commissioned agency to provide a 'health first' biopsychosocial case management approach for long-term IB recipients (of 3 years or more). Telephone and face-to-face techniques were used to address health needs (including behaviours) and any other related barriers to health or employment (such as debt or housing). The scheme was intended to complement mainstream services with case-managers signposting to NHS, Department for Work and Pensions (DWP) and other services. They also enrolled patients directly onto specially commissioned physiotherapy and counselling services. Patients were referred onto the programme by other NHS services (such as the Alcohol Service), their GPs, or they could self-refer (19.8%). The length of engagement with the service varied according to the needs of each service user (six months average). Participants were discharged when they were assessed to be ready to enter mainstream services such as Pathways to Work or community health services.

Non-Equivalent Comparison Group

A comparison group were recruited through IB 'Choices' events run by the regional Job Centre Plus. IB 'Choices' events offered a consistent sampling frame as all of those eligible within a given postcode area (IB receipt > 3 years) were invited to the event and attended on a voluntary basis. It is possible that those in the comparison group were therefore more motivated than the majority of the IB population. The comparison group received usual NHS care via general practice as well as usual access to Department for Work and Pensions activities. This included access to Pathways to Work, other vocational services and usual community health services (such as mental health services).

Data Collection

Socio-demographic (gender, age, housing tenure), social capital (contact with family and friends and participation with the wider community) and work history (previous jobs time spent in the job, time spent on IB) data were collected for both groups with questions taken from national surveys such as the General Household Survey (GHS). These data were collected to further understand the client group in relation to the regional and UK population norms and because such factors might impact on the effectiveness of the intervention (e.g. there is a strong relationship between occupational status and health outcomes, [13]). General health was measured via EQ5-D and EQ-VAS. General physical health was measured using SF-8 PCS (a shorter version of SF-36) and general mental health was measured using the SF-8 MCS. Two more condition specific measures - the Hospital Anxiety and Depression Scale (HADS), and the Nordic Musculoskeletal questionnaire - were also included as the two largest clinical categories of IB recipients in the UK are those with mental health or musculoskeletal issues. Data on health behaviours (tobacco and alcohol consumption) were also collected. (See Box 2).

Statistical Analysis

For this pilot study a hypothesis generating (rather than testing) approach was adopted, with the focus on inferences generated from the most parsimonious models for health outcomes. The repeated observations from individuals at the different time points imply inherent dependency in the data, which was accounted for through a fixed effect linear model with correlated errors (Repeated Measures ANOVA). The models assume time-related improvements of the health outcomes with either constant

or non-constant rates of improvement during both the intervention and post-intervention periods. We used models with linear trends (intercept and slope) for constant rate of improvement of the health outcomes and models with quadratic effects to capture non-constant rate of improvement with decline in health outcome post-intervention. Likelihood ratio test statistics were used to choose between models with linear or quadratic trends for each of the health outcome accounting for other factors at baseline - notably gender and age – that remained significant after adjusting for time, age and gender. We investigated for each health outcome, whether the correlation of the errors was constant or time dependent in order to choose an optimum covariance structure for each outcome. Based on the most parsimonious model for each health outcome, we investigated whether the rates of improvement in health of the intervention group over time (in months) differed from that of the comparison group accounting for other factors such as gender and age. In addition, we sensitized the models for deviation from Gaussian distribution assumptions by obtaining bootstrap-based confidence intervals. The effects of missing observations were investigated by using multiple imputations. Analyses for binary response (Yes/No) for Smoking and Drinking were analysed using Generalised Estimating Equations (GEE) whilst the number of smoking per day and units of drinks per week were analysed using fixed effects models with correlated errors. Descriptive statistics were analysed in SPSS® software, fixed effect linear models with correlated errors were performed using *`nlme'* package in R and the sensitivity for missing data was done using SAS/STAT® software.

Results

Participation

From October 2009 to March 2012, 459 participants received the case management intervention. Of these, 131 participants recruited between September 2009 and June 2010 were included in the evaluation of health outcomes. For the comparison group, 229 participants were recruited in the same period from 1429 attendees at IB 'Choices' events (16%). Baseline details for the intervention and comparison groups are shown in Table 1. The groups were statistically comparable in terms of gender, occupational class, time unemployed and smoking behaviour, but differed in age, marital status, use of social housing, primary health problem and use of alcohol. Intervention participants were more likely to have primarily mental health problems and worse health scores. However, using likelihood ratio test statistics to evaluate the importance of these baseline differences found no

significant association between them and health outcomes after adjusting for time, age and gender so they were not included in the final analysis. The completeness of follow-up data is shown in Table 2, with the effect upon findings of missing data assessed by imputation in relation to T2 where there was a low intervention group response (n=44). The number of participants was included in the analysis as they were to retain as much power as possible since our statistical approach allows the patients to contribute proportionally to their available data. The imbalance in numbers between the intervention and comparison groups constitutes more statistical power than forcing a balanced number of participants in both groups, which would have resulted in most of the observed data to be left unused.

Intervention effects

The comparative effect upon participants and non-participants is reported in Table 3. For each outcome a regression model provides age and gender adjusted estimates. Each model allows for linear (Time) and non-linear ($Time^2$) effects and for the incremental effect of the intervention as a constant (Int), linearly (Int.Time) and non-linearly ($Int.Time^2$) over time. Whilst the health of the comparison group remained stable, the general health (EQ5D, EQ-VAS) and the mental health (HADS-A, HADS-D, SF-8 MCS) of the intervention group improved, although this health improvement was less evident in terms of physical health (SF-8 PCS). There was no improvement in the musculoskeletal (Nordic 2) outcome (38% of participants in the intervention group had these problems as their primary health condition). There was no improvement in health behaviours in the intervention group and alcohol consumption actually increased.

The impact on EQ-5D of the intervention over time is shown in Figure 1. Variation in response at each time point is shown as error bars (95% confidence intervals) and the predicted (model) changes over time are shown as lines. There was a statistically significant difference at baseline between the intervention and the comparison group, with the intervention group reporting a poorer state of health: 0.30 (95%CI: 0.24 to 0.36) versus 0.42 (95%CI: 0.37 to 0.46). After six months, the EQ5D score of the intervention group had improved to a level similar to the comparison group, a gain persisting at 9 months (T4). The model-based inference assumes data distributions are Gaussian distribution, although bootstrap (non-parametric) methods produced comparable confidence intervals. The

missingness mechanism satisfied the assumption that incomplete data were missing at random (MAR) and thus incompleteness does not appear to influence the findings.

A similar pattern to EQ-5D was found for the EQ-VAS (Visual Analogue Scale) and the HADS-D scores. Unlike these measures, the HADS-A (figure 2) was best modelled by a linear improvement over time in the intervention group compared to the comparison group; the SF8 MCS was similar. No comparative changes were found for the Nordic-2 measure or SF8-PCS.

Cost-utility estimation

The case management intervention was delivered to a total of 459 participants at a cost of £1,161,047, or £2,530 per participant. Contemporary use of other healthcare and social resources used by the two groups was not recorded. Additionally, if the intervention had resulted in greater return to work this may have substantially offset the intervention costs, but this is not known. If baseline health outcomes were stable in the time preceding intervention then changes may be reasonably attributed to the intervention since no contemporaneous change occurred in the comparison group. Using trapezoidal estimation, the increase in quality of life (EQ-5D) shown in Figure 1 was estimated at 0.108 QALYs per participant within the study duration, providing a tentative cost-utility estimate of £23,500/QALY. This might be argued to be too conservative since there would be some continued benefit beyond the duration of follow-up. Extrapolating the model prediction to 14 months (when the curve returns to the baseline value) provides a higher estimate of 0.152 QALYs and cost-utility estimate of £16,700/QALY. These estimates are necessarily retrospective and approximate as the study was not specifically designed to test cost-utility. Subsequent research should capture other key resources - use of other health, social service, and return to work services.

Discussion

Main finding of this study

Those recruited to the case management intervention were initially in worse health than those in the comparison group, ascertained by generic (EQ-5D, EQ-VAS) and mental health scores (HADS-A, HADS-D, SF-8 MCS) and they exhibited worse health behaviours. By the end of the intervention their health had improved to levels of health similar to the comparison group. However both group's scores

remained well below the UK population norms for the selected health measures and both intervention and comparison group participants remained in receipt of IB. In contrast to generic and mental health improvements, impact upon physical health and musculoskeletal problems was limited with no improvements in the musculoskeletal (Nordic 2) outcome. This meant that the intervention was not effective in addressing the primary health condition of 38% of the intervention participants. These positive findings need to be understood in relation to the limitations of the study as detailed below.

The intervention was potentially cost-effective (in terms of the EQ-5D outcome) given current national policies for investment as the lower estimate of £16,700/QALY is below the threshold of £20,000/QALY given by NICE for case-management interventions that result in >=1% return to work rate [12]. A statement on whether the intervention is cost-effective though cannot be made until a comparative alternative study (e.g. longer-case management period) is undertaken. We also cannot state whether the estimate is generalisable to the whole IB population due to the methods of recruitment into the study including the proportion of self-referred patients who may be more motivated and thus more likely to experience a health improvement. These findings should be replicated using a more robust study design including use of randomisation to enhance attribution.

What is already known on this subject?

There is already considerable evidence about the successful use of case management approaches in health and social care, and in the provision of vocational support for those with long term health problems [12; 14; 15; 16]. However, no studies to date have evaluated the effectiveness of case management approaches for health improvement amongst people in long-term receipt of IB in England. As noted earlier, the need to find new and effective interventions to improve the health of those on long-term sickness benefits has been of increasing concern to successive UK governments [10; 11]. The intervention evaluated here builds on previous initiatives such as the Condition Management Programme (CMP). Whilst CMP was not regarded as wholly successful by the Department for Work and Pensions (DWP), the NHS found value in *"case management coupled with group and individual interaction"* [17: p28). Our evaluation also suggests that case management could have beneficial health effects for those in long term receipt of IB suggesting that CMP and related interventions could be useful components of the new DWP commissioned Work Programme.

However, whilst NICE evidence suggests that case management can be effective for both musculoskeletal and non-musculoskeletal conditions [12], within our study there appears to have been no improvement for those with musculoskeletal issues. This may have been because the service was not intensive enough, of sufficient duration or because the point of intervention (after 3 years on IB) was too late. There is evidence to suggest that musculoskeletal conditions require early intervention and that longer absence from work diminishes intervention effectiveness [18]. The association between musculoskeletal conditions and mental health also needs to be considered [19, 20, 21] with 43% of the intervention group reporting at baseline that they experienced both. So, those participants with a primary musculoskeletal condition may still have benefitted from the intervention via an improvement in a secondary - mental health - condition.

What this study adds

This is the first study to conduct a longitudinal and comparative evaluation of a health improvement case management approach for long-term IB recipients in England. The study shows the potential positive impact upon health of case management over a nine-month timescale including post-intervention follow-up and a cautious estimate of cost-utility for comparative purposes with future interventions.

Limitations of the study

The pilot study is limited to relatively small numbers of participants, voluntary participation in the intervention, the self-reported nature of the outcomes, low response at T2, and the non-randomised, non-equivalent study design. However, for a non-equivalent group design and key feature is not the comparability of the groups per se but how group measures change over time. The study also used validated health outcomes and provided a nine month follow-up. The QALY estimates are to be taken as with care due to fluctuating participation within the intervention group, the need to impute missing data values (particularly for T2) and the relatively small sample size. The analysis nonetheless indicates the scale of costs and utility to be expected in interventions of this nature.

Conclusion

This pilot study provides support for the health improvement benefits of case management approaches for IB recipients and tentative evidence of potential cost-utility when compared with NICE guidance on case-management of long-term incapacity for work [12]. Generic and mental health measures consistently improved in the intervention group, although there were no improvements in musculoskeletal outcomes or health behaviours. However, large scale, preferably randomised, research is required to examine the robustness and generalisability of these findings to other populations [22].

Research Ethics

This study received NHS National Research Ethics Service ethical approval from the County Durham and Tees Valley 2 Research Ethics Committee (REC reference 09/H0908/84).

Funding

This work was supported by County Durham and Darlington Primary Care Trust.

Author Contribution Statement

Principal investigator CB designed and oversaw all stages of the study with support from JW and co-applicants MB and JM. Data collection was conducted by JW and KG and data analysis was conducted by AK with input from MB and JM. JW and CB drafted the article with input from JM, AK, KG and MB.

Acknowledgements

We would also like to thank Graeme Greig, Sandra Moran, Mark Kennedy, Elaine Richardson, Carole Parker and Joanne Mounter. We also wish to thank all those who supported the project including Tim Blackman (former Director of the Wolfson Research Institute, Durham University) Rebecca Maier, with her help with ethical approval procedures, Jennifer Reynolds, Jacquelyn Briggs, Ladan Cockshut, Katie Ridley and Kerry Joyce for their help administering the project and carrying out telephone interviews.

References

1. Schuring, M., Burdorf, A., Kunst, A.E., and Mackenbach, J. (2007). The effect of ill health on entering and maintaining paid employment: evidence in European countries. *Journal of Epidemiology and Community Health*, 61, 597–604
2. Gabbay, M., Taylor, L., Sheppard, L., Hillage, J., Bambra, C., Ford, F. et al. (2011). NICE's Guidance on long term sickness and incapacity. *British Journal of General Practice*, 61: e118-124
3. Bambra, C. (2011) Work, worklessness and the political economy of health. Oxford University Press.
4. Bambra, C., Whitehead, M., and Hamilton, V. (2005) Does "Welfare to Work" work? A systematic review of the effectiveness of the UK's Welfare to Work programmes for people with a chronic illness or disability. *Social Science and Medicine,* 60, 9, pp. 1905-1918.
5. Bambra, C. (2006) The influence of government programmes and pilots on the employment of disabled workers. In: K. Needels and B. Schmitz, eds. *Economic and social costs and benefits to employers for retaining, recruiting and employing disabled people and/or people with health conditions or an injury: a review of the evidence*. Department for Work and Pensions Research Report no. 400. London: Department for Work and Pensions.
6. Clayton, S., Bambra, C., Gosling, R., Povall, S. Misso, K., and Whitehead, M. (2011) Assembling the evidence jigsaw: insights from a systematic review of UK studies of return to work initiatives for disabled and chronically ill people. BMC Public Health, 11:170.
7. Houston, D. S. & Lindsay, C. (2010): Fit for work? Challenges for the reform of disability benefits in the UK Policy Studies. 31, 2, p. 133-142.
8. Bambra, C. (2010) Doctors key to welfare reform. *British Medical Journal*, 341:c6029.
9. Warren J, Garthwaite K.A and Bambra C. (2013) 'A health problem ? Health and employability in the UK labour market', in *Disability Benefits and Employment Policy: Fit for Work, Fit for Purpose?* Eds Houston D and Lindsay C Palgrave Macmillan.
10. Black , C 2008 Working for a Healthier Tomorrow London, HMSO
11. Black C and Frost D 2011 Health at Work-an independent review of sickness absence. London, HMSO
12. NICE (National Institute for Health and Clinical Excellence). (2009). *Public Health Guidance 19: Managing long-term sickness absence and incapacity for work.* London: NICE
13. Marmot, M. (2010). Fair society, Healthy Lives: the Marmot review. London: University College
14. Burns, T., Catty, J., Becker, T., Drake, R.E., Fioritti, A., Knapp, M., Lauber, C., Rössler, W., Tomov, T., van Busschbach, J., White, S., and Wiersma, D. (2007) The effectiveness of supported employment for people with severe mental illness: a randomised controlled trial. *Lancet*, 370: 1146–52.
15. Kellett, S., Bickerstaffe, D., Purdie, F., Dyke, A., Filer, S., Lomax, V. and Tomlinson, H. (2011), The clinical and occupational effectiveness of condition management for Incapacity Benefit recipients. *British Journal of Clinical Psychology*, 50: 164–177.
16. Squires, H., Rick, J., Carroll, C., and Hillage, J. (2011) Cost-effectiveness of interventions to return employees to work following long-term sickness absence due to musculoskeletal disorders. *Journal of Public Health*, August 2011 fdr057v1-fdr057.
17. Department for work and Pensions (2012). The DWP funded and NHS delivered Condition Management Programme Lessons Learned. London, HMSO.
18. Squires, H., Rick, J., Carroll, C., & Hillage, J. (2012). Cost-effectiveness of interventions to return employees to work following long-term sickness absence due to musculoskeletal disorders. Journal of Public Health, 34(1), 115-124.
19. Parkes K R, Carnell S and Farmer E (2005) Musculo-skeletal disorders, mental health and the work environment HSE Research Report 316 HSE Books
20. Patten SB, Williams JV, Wang J: Mental disorders in a population sample with musculoskeletal disorders. BMC Musculoskelet Disord 2006, 25(7):37
21. Tüzün EH: Quality of life in chronic musculoskeletal pain. Best Pract Res Clin Rheumatol (2007), 21:567-579.
22. Skivington, K., McCartney, G., Thomson, H., and Bond, L. (2010) Challenges in evaluating Welfare to Work policy interventions: would an RCT design have been the answer to all our problems? BMC Public Health, 10:254.

Table 1: Baseline Characteristics of Survey Participants

	Intervention, N=131	Comparison, N=229	P
Gender			
Male	65 (50%)	115 (50%)	0.913
Female	66 (50%)	114 (50%)	
Age			
Mean and range, years	45 (21-64)	49 (19-63)	<0.001
≤ 45 years	57 (44%)	65 (29%)	
Marital Status			
Married	41 (31%)	107 (48%)	0.012
Divorced	32 (24%)	63 (28%)	
Single	45 (34%)	44 (19%)	
Tenure			
Renting	76 (58%)	132 (58%)	0.312
Renting (Social Housing)	49 (60%)	116 (85%)	<0.001
Transport			
No motor vehicle access	62 (47%)	97 (42%)	0.361
Occupational Class (last job)			
Professional	7 (6%)	7 (3%)	0.152
Intermediate	8 (6%)	20 (9%)	
Skilled Non Manual	8 (6%)	16 (7%)	
Skilled Manual	27 (21%)	33 (15%)	
Semi Skilled	27 (21%)	72 (32%)	
Unskilled	49 (39%)	74 (33%)	
Workless Households	97 (74%)	149 (65%)	0.078
Time spent on IB/ESA			
Mean (months)	98	108	0.170
Primary Health Problem			
Musculo-Skeletal	49 (38%)	110 (50%)	<0.001
Mental Health	62 (48%)	53 (24%)	
Digestive/Gastric	4 (3%)	23 (10%)	
Cardiovascular	7 (5%)	21 (10%)	
Respiratory	4 (3%)	5 (2%)	
Other	3 (2%)	10 (5%)	
Multiple (≥3) health problems	56 (43%)	130 (59%)	0.008
Seen clinician in past 30 days	107 (82%)	183 (80%)	0.684
Smoking and Drinking			
Regular Smokers	56 (43%)	83 (36%)	0.223
Cigarettes per day (in smokers)	54 (19%)	83 (19%)	1.000
Drinking Alcohol	80 (61%)	109 (48%)	<0.001
Units per week (in drinkers)	24.6(36.9)	18.6(38.2)	0.288
Health Scores			
EQ5D	0.30 (0.34)	0.42 (0.33)	0.002
EQ-VAS	42.08 (21.28)	46.45 (19.43)	0.048
SF8-PCS	34.24 (11.59)	33.24 (9.38)	0.372
SF8-MCS	33.72 (11.78)	36.86 (12.14)	0.018
HADS-A	12.68 (4.14)	10.54 (4.96)	<0.001
HADS-D	10.70 (4.36)	8.85 (4.44)	<0.001
NORDIC-2	3.40 (2.55)	4.30 (2.60)	0.002

Table 2 Invention and Comparison Group completeness of data

	Initial Questionnaire (T1)	Recall 1 (T2)	Recall 2 (T3)	Recall 3 (T4)
Intervention Group	131	44	79	95
Comparison Group	229	188	166	154

Table 3 Parameter estimates, standard errors, asymptotic and bootstrap confidence intervals for the health outcomes (p≤5% in bold)

Outcome	Parameter	Estimates	Std. Error	Asymptotic 95% Lower	Asymptotic 95% Upper	Bootstrap 95% CI Lower	Bootstrap 95% CI Upper
EQ5D	Intercept	0.728	0.093	0.546	0.910	0.545	0.890
	Gender	0.028	0.030	-0.031	0.087	-0.031	0.085
	Age	**-0.007**	**0.002**	**-0.010**	**-0.003**	**-0.010**	**-0.003**
	Int	**-0.148**	**0.037**	**-0.221**	**-0.076**	**-0.221**	**-0.077**
	Time	0.001	0.008	-0.015	0.018	-0.017	0.019
	$Time^2$	0.000	0.001	-0.002	0.002	-0.002	0.002
	Int.Time	**0.056**	**0.016**	**0.025**	**0.087**	**0.021**	**0.091**
	$Int.Time^2$	**-0.004**	**0.002**	**-0.007**	**-0.001**	**-0.008**	**0.000**
EQ-VAS	Intercept	53.099	5.530	42.261	63.937	41.322	63.533
	Gender	-2.307	1.770	-5.775	1.162	-5.977	1.194
	Age	-0.119	0.108	-0.331	0.093	-0.328	0.123
	Int	**-4.899**	**2.239**	**-9.288**	**-0.509**	**-9.372**	**-0.541**
	Time	-0.313	0.636	-1.561	0.934	-1.595	0.883
	$Time^2$	0.078	0.074	-0.067	0.222	-0.065	0.230
	Int.Time	**3.842**	**1.180**	**1.528**	**6.155**	**1.639**	**6.196**
	$Int.Time^2$	**-0.346**	**0.137**	**-0.614**	**-0.077**	**-0.632**	**-0.082**
HADS-A	Intercept	12.351	1.414	9.580	15.123	9.863	14.875
	Gender	0.388	0.460	-0.541	1.289	-0.506	1.298
	Age	-0.044	0.028	-0.098	0.011	-0.096	0.004
	Int	**1.852**	**0.531**	**0.811**	**2.894**	**0.845**	**2.893**
	Time	**-0.131**	**0.031**	**-0.191**	**-0.070**	**-0.188**	**-0.070**
	Int.Time	**-0.131**	**0.051**	**-0.232**	**-0.031**	**-0.250**	**-0.018**
HADS-D	Intercept	8.184	1.328	5.581	10.787	5.808	10.820
	Gender	0.514	0.433	-0.335	1.363	-0.337	1.327
	Age	0.008	0.026	-0.044	0.059	-0.045	0.055
	Int	**1.794**	**0.493**	**0.829**	**2.760**	**0.846**	**2.854**
	Time	0.038	0.093	-0.143	0.220	-0.124	0.200
	$Time^2$	-0.009	0.009	-0.027	0.009	-0.025	0.008
	Int.Time	**-0.606**	**0.180**	**-0.959**	**-0.254**	**-1.001**	**-0.198**
	$Int.Time^2$	**0.051**	**0.019**	**0.013**	**0.089**	**0.002**	**0.095**
NORDIC-2	Intercept	0.097	0.705	-1.286	1.479	-1.093	1.481
	Gender	**-0.465**	**0.229**	**-0.915**	**-0.016**	**-0.938**	**-0.050**
	Age	**0.090**	**0.014**	**0.062**	**0.117**	**0.064**	**0.112**
	Int	-0.502	0.264	-1.020	0.015	-1.038	-0.009
	Time	-0.002	0.016	-0.033	0.030	-0.032	0.031
	Int.Time	0.027	0.026	-0.024	0.079	-0.032	0.095
SF8-PCS	Intercept	47.294	2.885	41.640	52.949	40.744	53.380
	Gender	**2.310**	**0.939**	**0.468**	**4.151**	**0.356**	**4.312**
	Age	**-0.309**	**0.057**	**-0.420**	**-0.198**	**-0.432**	**-0.187**
	Int	-0.286	1.086	-2.415	1.844	-2.368	1.910
	Time	0.171	0.240	-0.299	0.641	-0.250	0.570
	$Time^2$	-0.014	0.026	-0.064	0.036	-0.057	0.032
	Int.Time	**1.678**	**0.432**	**0.830**	**2.525**	**0.736**	**2.547**
	$Int.Time^2$	**-0.162**	**0.047**	**-0.255**	**-0.070**	**-0.262**	**-0.062**
Sf8-MCS	Intercept	33.277	3.378	26.656	39.897	27.130	40.900
	Gender	-0.610	1.091	-2.748	1.529	-2.730	1.444
	Age	0.087	0.066	-0.043	0.217	-0.059	0.205
	Int	**-3.032**	**1.303**	**-5.585**	**-0.478**	**-5.781**	**-0.446**
	Time	**0.259**	**0.100**	**0.063**	**0.454**	**0.067**	**0.455**
	Int.Time	**0.404**	**0.157**	**0.096**	**0.713**	**0.060**	**0.736**

Smoking	Intercept	0.743	0.622	-0.476	1.963	-	-
	Gender	0.303	0.209	-0.107	0.712	-	-
	Age	**-0.030**	**0.0123**	**-0.054**	**-0.006**	-	-
	Int	0.185	0.228	-0.262	0.634	-	-
	Time	0.013	0.009	-0.005	0.030	-	-
	Int.Time	-0.017	0.018	-0.052	0.018	-	-
Smoking /day	Intercept	11.559	5.099	1.484	21.634	-	-
	Gender	3.438	1.810	-0.138	7.014	-	-
	Age	0.096	0.101	-0.104	0.296	-	-
	Int	1.435	2.087	-2.689	5.560	-	-
	Time	-0.285	0.155	-0.590	0.019	-	-
	Int.Time	0.111	0.240	-0.362	0.583	-	-
Drinking	Intercept	0.122	0.596	-1.046	1.290	-	-
	Gender	**0.637**	**0.191**	**0.264**	**1.011**	-	-
	Age	-0.011	0.012	-0.034	0.012	-	-
	Int	**0.562**	**0.228**	**0.115**	**1.008**	-	-
	Time	**0.055**	**0.015**	**0.025**	**0.084**	-	-
	Int.Time	**-0.080**	**0.027**	**-0.134**	**-0.027**	-	-
Units/week	Intercept	33.933	11.192	11.882	55.984	-	-
	Gender	**8.445**	**3.699**	**1.157**	**15.733**	-	-
	Age	**-0.492**	**0.216**	**-0.917**	**-0.066**	-	-
	Int	6.082	4.176	-2.146	14.310	-	-
	Time	-0.469	0.250	-0.961	0.024	-	-
	Int.Time	-0.466	0.409	-1.271	0.339	-	-

Figure 1: Average possible health levels of EQ-5D for the intervention and comparison groups. The error bars represent the observed profile and the fitted lines represent the predicted profiles.

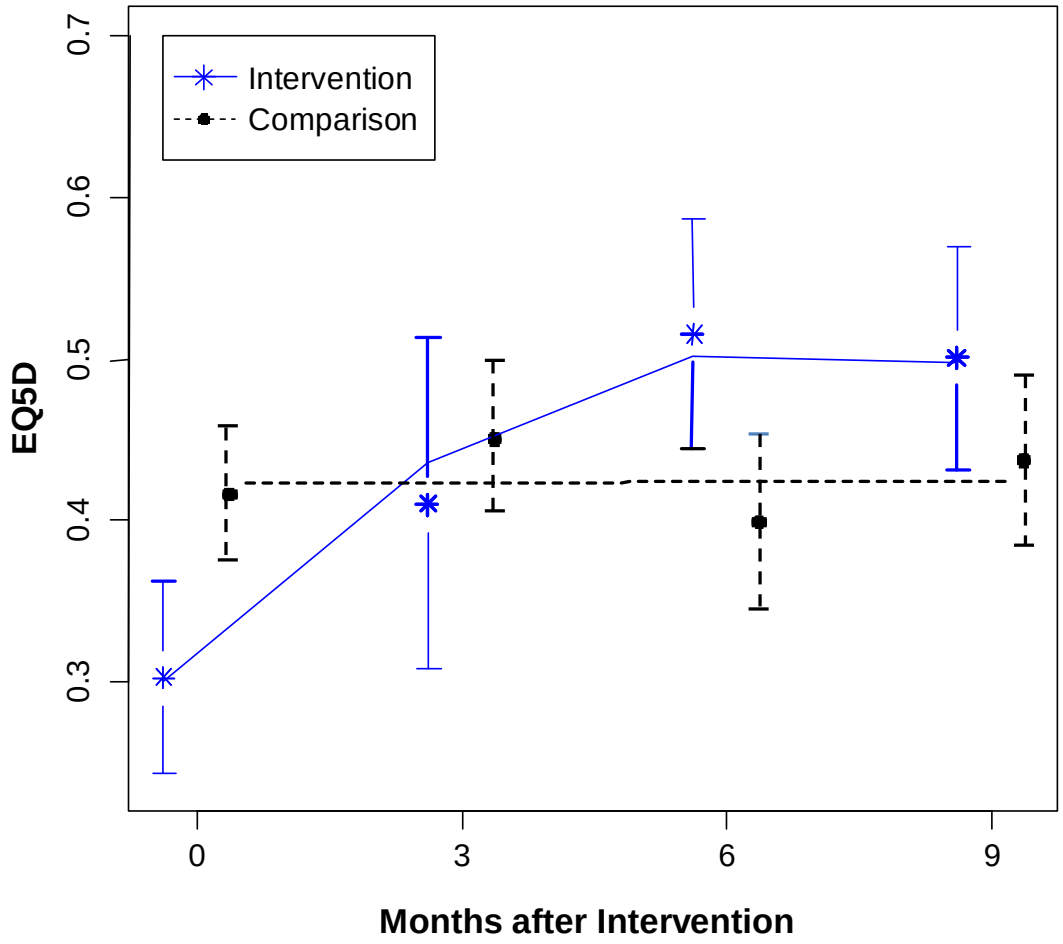
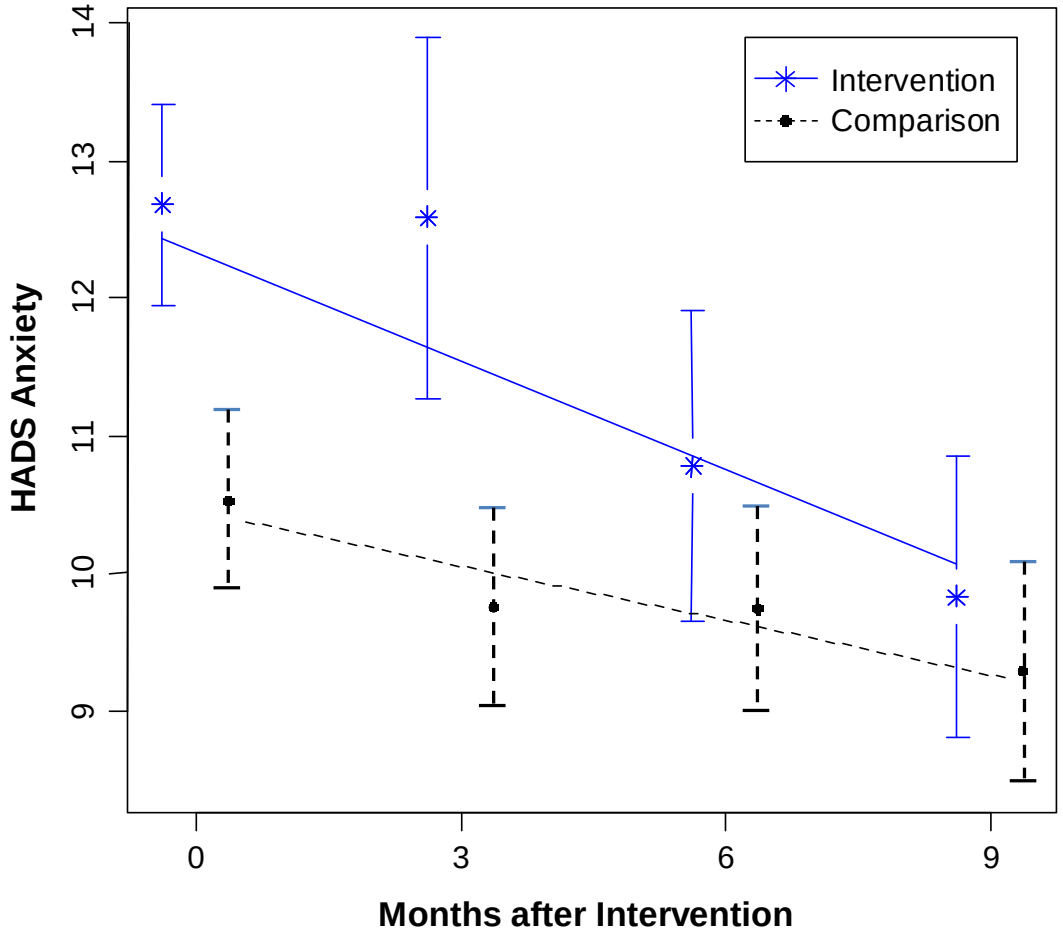


Figure 2: Average possible health levels of HADS-A for the intervention and comparison groups. The error bars represent the observed profile and the fitted lines represent the predicted profiles.



Box 1: **National Institute for Health and Clinical Excellence recommendation on return to work interventions for Incapacity Benefit recipients [2, 10]**

Who is the target population?

People with health problems who are unemployed and claiming Incapacity Benefit or Employment Support Allowance.

Who should take action?

Department for Work and Pensions and other bodies or organisations which may commission services for those who are unemployed and claiming Incapacity Benefit or Employment Support Allowance.

What action should they take?

Commission an integrated programme to help claimants enter or return to work (paid or unpaid). The programme should include a combination of interventions such as: an interview with a trained adviser to discuss the help they need to return to work; vocational training, including that offered by *New Deal for Disabled People* (for example, help producing a curriculum vitae, interview training and help to find a job or a workplacement); a *condition management* component run by local health providers to help people manage their health condition; financial measures to motivate them to return to work (such as return-to-work credit); support before and after returning to work (this may include one or more of the following: mentoring, a job coach, occupational health support or financial advice).

Box 2: **Detailed description of health measures**

EuroQol (EQ-5D) and (EQ-VAS)
Two parts: a questionnaire and a 'health thermometer'. The EQ-5D questionnaire asks participants about their mobility, ability to self-care, their ability to carry out their usual activities, pain and discomfort and anxiety and depression on the day when they are interviewed. The responses are converted to a value between 0 and 1. The higher the value is the better the health state. The second element is the Visual Analogue Scale, often known as a 'Health Thermometer'. Participants are asked to rate their health on the day they are interviewed on a scale of 0 -100. 0 represents the worst health state the participant can imagine, 100 represents the best health state they can imagine with 50 representing the midpoint.

Hospital Anxiety and Depression Scale (HADS)
There are two parts: HADS-A (Anxiety) and HADS-D (Depression). Both ask participants to choose options that best describe how they are feeling. Both generate a score between 0-21. A higher score indicates a higher degree of Depression.

Quality Metric Short Form 8 (SF8)
SF-8 is a measure of health that produces a physical health score (PCS) and a mental health score (MCS). Participants are asked 8 questions about their health during the past four weeks. These generate two scores, both between 0-100: the higher the score the better the health state.

Nordic Musculoskeletal Questionnaire (element 2)
There are three elements of which only the second element (Nordic-2) was appropriate to this study Nordic-2 is a measure of musculoskeletal problems over the preceding 7 days. Participants are asked whether they have had problems with different areas of the body. The measures produce a scores of between 0 (no problem) areas and 9 (nine problem areas).