Effectiveness and costs of a vocational advice service to improve work outcomes in patients with musculoskeletal pain in primary care

Wynne-Jones, G.; Artus, Majid; Bishop, Annette; Lawton, Sarah L; Lewis, M; Jowett, Sue; Kigozi, Jesse; Main, C; Sowden, Gail; Wathall, S.; Burton, K; van der Windt, Danielle A W M; Hay, Elaine M; Foster, N. E.

Document Version
Peer reviewed version

Link to publication on Research at Birmingham portal

Effectiveness and costs of a vocational advice service to improve work outcomes in patients with musculoskeletal pain in primary care: A cluster randomised trial (SWAP trial ISRCTN 52269669)

Gwenllian Wynne-Jones[1]*, Majid Artus[1], Annette Bishop[1], Sarah A Lawton[2], Martyn Lewis [1,2], Sue Jowett[1,3], Jesse Kigozi[1,3], Chris Main[1], Gail Sowden[1,4], Simon Wathall[1,2], A Kim Burton [5], Danielle A van der Windt[1], Elaine M Hay[1], Nadine E Foster[1,2] and the SWAP study team

*Corresponding author

[1]Arthritis Research UK Primary Care Centre

Research Institute of Primary Care and Health Sciences

Faculty of Medicine and Health Sciences

Keele University, ST5 5BG

Tel: +44 (0) 1782 733962

Fax: + 44 (0) 1782 733911

Email: g.wynne-jones@keele.ac.uk

[2]Keele Clinical Trials Unit

Keele University

Staffordshire

ST5 5BG

[3]Health Economics Unit

Institute of Applied Health Research

College of Medical and Dental Sciences,

University of Birmingham

Edgbaston

Birmingham, B15 2TT

[4]IMPACT Service,

Haywood Hospital,

High lane,

Burslem,

Stoke-on-Trent, ST6 7AG

[5]Centre for Applied Research in Health

University of Huddersfield

Queensgate

Huddersfield, HD1 3DH

Number of text pages: 28

Number of figures: 2 plus 1 online figure

Number of tables: 5 plus 2 online tables

Abstract: Musculoskeletal pain is a common cause of work absence and early intervention is advocated to prevent the adverse health and economic consequences of longer-term absence. This cluster randomised controlled trial investigated the effect of introducing a vocational advice service, into primary care to provide occupational support. Six general practices were randomised, patients were eligible if they were consulting their general practitioner (GP) with musculoskeletal pain, were employed and struggling at work or absent from work <6 months. Practices in the intervention arm could refer patients to a vocational advisor embedded within the practice providing a case managed stepwise intervention addressing obstacles to working. The primary outcome was number of days off work, over 4 months. Participants in the intervention arm (n=158) had fewer days work absence compared to the control arm (n=180) (mean 9.3 (SD 21·7) versus 14·4 (SD 27·7)) days, Incidence Rate Ratio (IRR) 0·51 (95% Confidence Interval 0·26, 0·99), p=0·048). The net societal benefit of the intervention compared with best care was £733: £748 gain (work absence) versus £15 loss (health care costs). The addition of a vocational advice service to best current primary care for patients consulting with musculoskeletal pain led to reduced absence and cost savings for society. If a similar early intervention to the one tested in this trial was implemented widely, it could potentially reduce days absent over 12 months by 16%, equating to an overall societal cost-saving of about £500 million (US $6 billion), and requiring an investment of only £10 million.

Keywords: Cluster randomised controlled trial; Vocational advice; Occupational advice; musculoskeletal pain; Primary care

Introduction

Musculoskeletal pain is one of the most common causes of work absence [1,2]. Across Europe almost a quarter of workers will experience pain in their neck, shoulders or upper limbs, and an estimated half of the European workforce will experience back pain at some point in their lives at a cost of approximately €12 billion overall [3]. The cost of work absence attributed to musculoskeletal pain in European Union countries is between 0.5 and 2.0% of national gross domestic product [3]; pain also has a considerable impact on individuals' earnings and associated costs to the state in benefit payments [4]. In the UK the estimated costs in 2003 for GP consultations only as a result of musculoskeletal conditions was £1.34 million [5]. The prevalence and incidence of many musculoskeletal conditions increase with older age; this, coupled with the rising retirement age, means that the impact of musculoskeletal pain on the workforce will rise further [3].

Remaining active at work, despite pain, has been demonstrated to be beneficial to individuals and employers resulting in less sickness absence, less time on modified duties and a reduction in pain recurrence [4]. Intervening early when employees report musculoskeletal pain can have a significant impact on their ability to remain in work [6,7]. However, the provision of independent occupational health services is scarce, and for the majority of working age people the first port of call for advice is their General Practitioner (GP) [6]. In the United Kingdom the GP is also the gatekeeper to health related benefits through the "Fit Note" system whereby absence of greater than seven days is sanctioned. However, many GPs report that they feel ill-equipped to manage occupational health issues and have had little or no training in the use of Fit Notes [8]. Previous initiatives to address health and work, have been aimed primarily at those with longer term absence [9,10]. However, given the evidence that

the longer an individual is out of work the less likely it is that they will return, intervening before an individual experiences long term absence may be beneficial to both the individual and wider society.

Primary care is likely to be the ideal setting in which to offer patients early access to appropriate occupational health support, also termed vocational advice, occupational advice, and workplace coaching in the literature. Whilst there are guidelines in place to support primary care practitioners in providing appropriate advice and support about work, implementation of these is variable [11]. Improvement in training and education about managing occupational health in primary care should be coupled with provision of services to which patients may be referred for advice and assistance about work.

The aim of this cluster randomised controlled trial was to determine whether the addition of a vocational advice service to best current primary care can reduce work absence in patients consulting their GP for musculoskeletal pain who are either absent from work or struggling to remain in work because of their pain.

Methods

Study design and participants

The methods are reported in full in the published protocol [12]. The Study of Work and Pain (SWAP) was a pragmatic, cluster randomised controlled trial in primary care with two parallel arms, an economic evaluation, and linked qualitative interviews (reported separately). The unit of randomisation was the general practice with data collected from individual participants.

Consenting GP practices were randomly assigned to provide either best current primary care for managing the impact of musculoskeletal conditions on work or the same best care plus the addition of a vocational advice service, located in the practice and staffed by trained Vocational Advisors (VAs) who provided occupational advice about working with musculoskeletal pain. GP practices were eligible for participation if they were located in the National Institute of Health Research Clinical Research Network: West Midlands (NIHR CRN: WM), which supports delivery of research within primary care practices in the region. Recruitment took place between 2012 and 2014 and participants were followed-up 4 and 12 months later. Patients were eligible for participation if they: were consulting with musculoskeletal pain; aged 18 to 70 years; currently in paid employment; had current sickness absence due to musculoskeletal pain of less than 6 months duration (either GP or self-certified and either first consultation for a Fit Note or a repeat consultation for a Fit Note), or were considered by the GP or Nurse Practitioner (NP) to be struggling with work due to musculoskeletal pain. Patients were not eligible for participation if they met any of the following criteria (full criteria are reported in the protocol) [12]: Patients with symptoms indicative of possible serious pathology, requiring urgent medical attention; those who have long term work absence (greater than 6 months); those with serious mental health problems. Eligible patients were identified when they consulted their GP/NP and were introduced to the study and given an information pack. The pack contained a letter of invitation, participant information sheet, consent form to participate in the research evaluation of the service, self-completion questionnaire and a pre-paid reply envelope. Eligible patients, not identified during the consultation, were later identified by the NIHR CRN: WM through regular medical record reviews (see published protocol for full details) [12]. Selection bias was

minimised in this cluster trial through identical methods of participant identification, invitation and recruitment at both intervention and control practices.

A trial steering committee and independent Data Monitoring Committee oversaw the trial. The National Research Ethics Service West Midlands – Staffordshire in the UK approved the protocol (REC reference: 12/WM/0020), and the trial was registered at ISRCTN 52269669.

Randomisation and masking

GP practices were the unit of randomisation. Practices were matched on registered population list size, the matched practices were randomly allocated to the intervention or control arms by stratified block randomisation. The randomisation process within the individual blocks was computer-generated by the trial statistician. GPs, NPs and VAs could not be blinded to allocation. Individual participants were informed that local musculoskeletal services were being evaluated and their consent was sought to participate in data collection and medical record review. The data were analysed independently by two statisticians one of whom was blinded to intervention allocation.

Interventions

Both intervention and control practices provided best current work-focussed primary care. The provision of best current care was supported by providing GPs and NPs with an education session lasting one hour. This emphasised four key messages: 1) work is usually good for people with musculoskeletal pain, 2) long periods of absence are generally harmful, 3) musculoskeletal pain can generally be accommodated at work, and 4) planning and supporting return-to-work are important aspects of clinical management [4, 13].

The intervention practices also hosted a new vocational advice service [12], and GPs and NPs could refer patients to the service whether, or not, patients consented to take part in the research evaluation. Patients referred to the VA were contacted 5 working days after referral. Initial contact was by telephone (step 1), with one or more face-to-face meetings (step 2) and contact with employers (step 3) being held subsequently, if required. VAs used the "psychosocial flags framework"[13] to assist patients in identifying and overcoming obstacles to returning to or remaining in work with their musculoskeletal pain. The VAs focussed discussions on three main areas; i) Psychological or behavioural obstacles to working e.g. beliefs about pain, illness behaviours (yellow flags) [14]; ii) Work perceptions e.g. the beliefs about the physical and social impact of work on health (blue flags) [15]; iii) Context factors e.g. objective working conditions and characteristics, and financial impact of working status such as job security and benefit entitlements (black flags) [13]. The VA and patient jointly developed a plan to manage health and work issues and to support the patient in addressing identified obstacles, with regular review. The VA also ensured that the patient's GP was included in communications using the practice communications system linked to the patient's medical record. This ensured that clinical issues identified as obstacles to work could be communicated to the GP for resolution, and that return to work plans could also be provided to the GP. Four healthcare practitioners were recruited to the VA role to deliver the service; three physiotherapists and one nurse (all vocational advice was actually delivered by the three physiotherapists), all completed a four day training course (developed by the study team and reported separately) and half day update prior to the start of the service. The VAs were new recruits to this role and did not provide any other services to the general practice. The service was "low intensity" and based on the principles of case management using a stepped care model to develop a goal orientated approach to remaining in or returning to work (Figure 1), along with the intention of getting the key players (person; healthcare;

workplace) onside [13]. Patients continued to be eligible for vocational advice until they achieved a sustained return-to-work (the patient returns to work and does not initiate contact with the VA for a period of at least 2 weeks) and felt able to manage their musculoskeletal pain in the context of their work, or until they had been absent from the workplace for a total of six months and qualified for Employment and Support Allowance.

Outcomes

Demographic data, health and work data were collected after GP consultation and, in the intervention practices, before an appointment with the VA, and at 4 and 12 months follow-up. Full details of the primary and secondary outcomes collected are provided in the protocol [12].

The primary outcome measure was number of days off work over 4 months, measured at the individual participant level. Work absence was identified at follow-up based on the following self-reported questions; "Have you taken time off work during the last 4 months (since your last questionnaire) because of your pain?", "If yes, please write the number of days, weeks or months you were off work due to your pain in the last 4 months". 'Days off work' in this context captures periods of self-certified absence as well as GP certified absence. For the purposes of this trial 1 week was classified as 5 days and 1 month as 21 days. Further analysis of time off work examined any self-reported time off work (binary yes/no) and GP certified periods over 12 months follow-up identified from the medical record. Secondary outcome measures included pain intensity (0-10 numerical rating scale), bothersomeness (1-5 rating scale) [16], global assessment of change (5 point rating of general health from excellent to poor), self-efficacy to return-to-work (Self-efficacy to Return-to-Work

Questionnaire) [17], work presenteeism (Stanford Presenteeism Scale 6)[18] and self-rated work performance (0-10 numerical rating scale).

Statistical analysis

For the primary outcome (days off work over 4 months), the analysis was by hierarchical negative binomial regression adjusting for age, gender, and GP practice size (at the GP-cluster level) [19]. The best-fitting model according to goodness-of-fit (higher log-likelihood, and lower Akaike's Information Criterion (AIC) and Bayesian Information Criteria (BIC)) was given by a zero-inflated model; hence, the hierarchical zero-inflated negative binomial (ZINB) regression was used for the analysis of time off work over 4 months (primary) and 12 months (secondary). Given the limited number of GP practices, the hierarchical model included individual practitioners (GPs and NPs) at the cluster-level; differences in GP behaviours are known to be a major influence in varying sickness certification prescribing practice [20]. Longitudinal mixed-models (linear- or generalised- as appropriate to numerical and categorical outcome data, respectively) were fitted to estimate and test for between-group effects across other outcome measures, adjusting for baseline covariates (age, gender and GP practice size). An intention-to-treat analysis was followed. The statistical analysis followed the plans described in the published protocol [12] and the final version of the Statistical Analysis Plan (SAP) agreed with the Data Monitoring Committee.

Sensitivity analyses

1. Evaluation of the primary outcome measure (number of days off work by robust Poisson and zero-inflated models)

2. Evaluation of the primary outcome measure (number of days off work) by a ZINB model with robust variance estimator [21] adjusted for (i) age, gender and practice size, and (ii) adjusted for age, gender, practice size plus baseline pain scores and days off work over the past 12 months..

3. Evaluation of the primary outcome measure utilising the GP practice as the unit of clustering rather than the individual GP/NP practitioner; including GP practice as a random factor intercept in the hierarchical model.

4. A per protocol evaluation (and complier average causal effect (CACE) evaluation) comparing time off work for those participants in the intervention practices who engaged with any aspect of the vocational advice service (at least one contact with a VA) versus (i) all control arm participants, (ii) 'comparable' participants in the control practices that would be expected to similarly adhere with treatment protocol – via an instrumental variable analysis (adherence / non-adherence, defined as at least one contact with the VA).

Subgroup analyses

Exploratory evaluation of the primary outcome was carried out to examine whether time off work appeared differed between subgroups. The three subgroup analyses agreed and documented in the SAP were: baseline return-to-work self-efficacy, location of pain (spinal pain versus pain in other areas), and duration of work absence (at least 10 days versus /less than 10 days). Statistical estimates were obtained through including interaction terms in the statistical model of treatment effect.

Sample size

The sample size calculation was based on the ability to detect a between group difference of at least 10 days off work at 4 months, given an expected standard deviation of 25 days [22], 80% power, and 5% two-tailed significance level. The sample size takes into account: (i) 30% inflation through clustering of data (at practitioner-level) based on an ICC for between-practitioner effects of 0.05 [23], variation in expected VA service referral rates between GPs (based on an expected coefficient of variation of 0.65) [24], and (ii) 25% inflation through allowance for 20% loss to follow-up at 4 months. This resulted in a required sample size of 330 participants (165 per arm).

Economic evaluation

An incremental cost-effectiveness analysis was undertaken using mean days off work as the measure of outcome, to calculate the cost per sick day avoided, from a healthcare perspective. Patient-level healthcare costs concentrated on National Health Service (NHS) and private healthcare resource use for musculoskeletal pain obtained from patient questionnaires at 4 and 12 months, and additional costs of the VA service (eTable 1, available online at http://links.lww.com/PAIN/A489). Hierarchical modelling was used to estimate differential costs and differential Quality Adjusted Life Years (QALYs) controlling for treatment arm and clustering [25]. Details of contact with the VAs were obtained through case report forms. Unit cost data relating to resource use are reported in eTable 2 (available online at http://links.lww.com/PAIN/A489), and a price year of 2013 was used, with costs presented in UK pounds (£). A cost-benefit approach was used to generate a net societal benefit and return-on-investment of using the VA service. Wider societal costs in relation to the VA intervention were assigned to self-reported work absence using the human capital approach by multiplying days off work during follow-up by the Standard Occupational Classification

(2010 edition) related respondent-specific wage rates. Discounting was not performed because of the 12 month follow-up period.

Public and Patient Involvement and Engagement (PPIE)

Patients with musculoskeletal pain and primary care clinicians involved in their treatment were involved throughout the SWAP trial, and were independent from those participating in the trial. PPIE representatives were involved in the development of the research question, were active members of the grant application with additional members involved in the trial steering committee and providing advice on all aspects of the design, recruitment and retention methods, as well as reviewing all patient facing materials.

Results

Recruitment

Twenty general practices were approached with six general practices being eligible; they were randomised, 3 to the intervention and 3 to the control arm. Participants were recruited between July 2012 and January 2014; Figure 2 shows the flow of participants through the trial. A total of 338 participants consented to participate in the research data collection after their consultation at participating practices, 158 to the intervention and 180 to the control arm. Follow-up was 75% (*n*=119) and 69% (*n*=109) at 4 and 12 months respectively in the intervention arm and 82% (*n*=148) and 73% (*n*=131) at 4 and 12 months in the control arm.

Baseline characteristics

Table 1 reports the baseline characteristics of participants, which were comparable. The mean age was 49.5 and 47.9 years, with 56% and 59% female in the intervention and control arms, respectively. The majority of participants were working full-time. Participants in the control arm reported that they had marginally more days of work absence in the previous 12 months.

At baseline, duration of symptoms, measures of pain intensity and bothersomeness were similar in both arms.

Adherence with treatment protocol

Of the 158 participants in the intervention practices, 120 (76%) were referred to the VA service (Figure 2). Of these, 97 (81%) had at least one contact with a VA. The average number of contacts between the VAs and patients was two with the majority of these being telephone contacts (89%) lasting an average of 13.3 minutes (Table 2). Exploration of health and work issues were frequently recorded by the VAs on case report forms, but return-to-work planning was not commonly recorded.

Primary outcome

4 months

At 4 months there was some evidence for effect in the number of days off work between arms with the intervention arm reporting fewer days off work mean of 9.3 (SD 21.7) days compared to 14.4 (SD 27.7) days in the control arm, an adjusted incidence rate ratio of 0.51 (p=0.048). Results of the sensitivity analyses including different model estimation, non-parametric testing, per-protocol/CACE-complier evaluation, and accounting for clustering at GP practice level concurred with the primary analysis in showing greater time (days) off work in the control arm (Table 3). The difference in days off work was largely accounted for by the lower number of GP certified days in the intervention arm at 8.4 (SD 21.0) days versus 13.5 (27.5) days in the control arm (p=0.020) (Table 3).

12 months

By 12 months there was no overall statistically significant difference in the cumulated number of days of work absence between arms. However the intervention arm reported fewer days off work certified by the GP at a mean of 16.4 (SD 34.2) days compared to 22.9 (SD 50.5) days in the control arm (p=0.018). The control arm reported fewer days self-certified than the intervention arm at a mean of 1.5 (SD 3.3) days compared to 3.9 (SD 15.0) days (p=0.001) (Table 3).

Exploratory subgroup analyses

At 12 months, exploratory subgroup analyses showed that the VA service was significantly more successful in those with spinal pain compared to those with other musculoskeletal pain (Incidence Rate Ratio (IRR) 0.25 (95% Confidence Interval (CI) 0.10, 0.62) ($p_{interaction}$=0.003). The intervention was also significantly more successful in those who had work absence that exceeded 10 days at baseline compared to those with absence periods of less than 10 days (IRR 0.30 (95% CI 0.11, 0.83) ($p_{interaction}$=0.020) (Table 3)). Baseline level of self-efficacy to return-to-work had little impact on the effect of the intervention (Table 3).

Secondary outcomes

Self-reported time off work (binary yes/no) was examined as a secondary outcome. Separate analysis compared the proportions of participants in the two trial arms issued with a GP certified fit note, assessed through medical records (Table 3). Of the health-related (secondary) outcome measures there were few statistically significant differences between the intervention and control arms for measures of pain, bothersomeness, pain self-efficacy, IPQ-R, HADS anxiety and depression and general health. Though estimated differences were small, the health outcomes were generally in favour of the intervention arm (Table 4). Work-

related measures demonstrated statistically significant differences between arms, in favour of the intervention arm, at both 4 and 12 months in return-to-work self-efficacy and performance at work, and a significant difference in presenteeism at 4 months (Table 4).

Economic evaluation

The VA service resulted in greater mean benefits in terms of days off work (6.7 fewer days off work; adjusted difference in time off work over 12 months), at slightly higher NHS and healthcare costs (cost difference of £48 and £15 for NHS and healthcare perspectives respectively) (Table 5). From an NHS perspective, this resulted in an incremental cost-effectiveness ratio of £7.20 per day of absence avoided.

The net societal benefit of the addition of the VA service compared with best current care alone was £733 (£748 gain (work absenteeism) minus £15 loss (healthcare-related costs)) demonstrating that the intervention represents more efficient use of resources than the control (Table 5). The corresponding return-on-investment (ROI) from a societal perspective was £49 (£733 divided by £15) – that is, every £1 invested in the VA service will return an estimated £49 ($64USD). The inclusion of training costs and monthly mentoring brings the ROI to £25 ($30USD).

The point estimate suggests that the intervention was more effective (with fewer days off work) and associated with higher costs than the control. eFigure 1 shows that for a willingness to pay of £40 per sick day avoided, the probability that the intervention is cost-effective was slightly over 50% (available online at http://links.lww.com/PAIN/A489).

Discussion

The SWAP trial demonstrated that the addition of a low intensity, early access, vocational advice service to best current primary care for adults consulting with musculoskeletal pain led to fewer days off work over 4 months, indicating some evidence for effect of the intervention. The intervention improved measures of work performance, presenteeism and self-efficacy to return-to-work. Use of the vocational advice service for musculoskeletal pain was associated with slightly higher costs but the cost-benefit analysis demonstrated the broader societal value of the VA service.

Implications

The VA service, also termed occupational advice and workplace coaching, highlighted two key implications relating to the study population and the intensity of intervention delivered.

Timing of the intervention

The sample included in the SWAP trial could be considered early in their "work absence career"; patients were eligible if they were struggling at work as well as those having a short period of absence (less than 6 months). Whilst the addition of the VA service led to significantly fewer days off work, exploratory subgroup analysis in those participants with <10 days absence versus ≥10 days but <6 months absent at baseline found that the intervention was more successful in those with the longer absence duration. Whilst early intervention is advocated [12] these results suggest that a VA intervention might be better targeted to those with more than 10 days (2 working weeks) of absence. van Duijn *et al* [26] reviewed the literature around timing of interventions for individuals on sick leave due to back pain, reporting the optimal window in which to intervene as 8 to 12 weeks. These

findings suggest the optimum time to provide support in managing health and work is likely to be after 10 days (approximately 2 working weeks) of absence, but this needs testing in future studies.

Intensity of intervention delivered

The intervention provided in the SWAP trial was low intensity with the majority of vocational advice delivered by telephone. This is in keeping with robust evidence that telephone based vocational advice can help a substantial proportion of cases to self-manage their health problem and may also facilitate return-to-work [27]. There is evidence that simple, low intensity interventions provide similar benefits to complex, multi-modal, interventions whilst avoiding unnecessary medicalisation. This is particularly pertinent to the SWAP trial where participants had short-term or no work absence and were in an ideal position to manage their condition with appropriate advice before their absence became long-term. The model of stepped care evaluated in the SWAP trial is similar to that proposed by Burton *et al* [27], requiring only those with more complex needs to access costly face-to-face contact.

Strengths and limitations

The SWAP trial has a number of strengths. It is the first trial to evaluate a VA service embedded in general practice offering biopsychosocial advice to people with musculoskeletal pain, a leading cause of work absence. The VA service was also acceptable to patients with 75% (253 patients) of those offered a referral accepting this offer. The SWAP trial is also the first to intervene so early including those who were struggling at work, with the aim of preventing future work absences. Whilst the stepped care vocational advice service was brief and mainly provided over the telephone, this method is supported by the literature showing

that brief vocational advice interventions are as effective as effort-intensive interventions [28] and there are robust data to support telephone-based interventions [27]. The Department for Work and Pension's evaluation of the Fit For Work service pilots [29] also found that low cost interventions (equating to low intensity interventions) were more likely to be the most cost effective, and many of these interventions included populations with longer term absence indicating that there would be utility in evaluating a similar VA service in those with longer absence duration. A further strength of this trial concerned activities to ensure continued engagement with general practices. This included a range of measures for both the intervention and control practices comprising; provision of an education session around managing health and work before the trial commenced; regular contact with the trial team GP; a GP "champion" in each practice who was the point of contact for the trial. In intervention practices, VAs actively engaged in practice life, joining breaks and staff meetings and providing both formal and informal feedback about the service to GPs. This was important given the difficulty in engaging GPs in studies of vocational advice and has been reported by Rannard *et al* (2014) [30] and the Fit for Work Pilots [29]. The finding that there was a difference in GP certified periods of absence could have been related to the visibility of the VAs in the practice, suggesting that raising the profile of available vocational advice services providing vocational advice may be of benefit. The qualitative analyses conducted alongside this trial was unable to elucidate the reasons for the decrease in the issue of fit notes [31] and further work is needed to identify whether the availability of a vocational advice service does change GP behaviour reducing the issue of fit notes or whether accessing vocational advice changes patients' behaviour in asking for certified absence.

There are several limitations. Firstly, the association between the intervention and the measures of work outcomes (return-to-work self-efficacy, performance and presenteeism) were influenced by the adjustment of practice size due to the small number of practices, 3 intervention and 3 control (practice size was adjusted for as it was the only stratification variable used in randomisation). Secondly, whilst three steps were available to the VAs in the delivery of the VA service, only one workplace visit was undertaken (step three), the reasons for which need some consideration. The VAs within the SWAP trial reported that participants were unwilling for them to contact their employers. Many participants were very early in their work absence and some were not currently absent, but struggling at work; the lack of employer visits may reflect the trial population and the primary care setting, where contact between vocational advisors and employers is uncommon, this is a finding in other similar studies [32]. A linked issue relates to the lack of recorded return-to-work plans on the case report forms of patients accessing the VA service in the intervention practices, this may be explained by the early nature of the participants' work absence. Whilst many participants received at least one phone call from the VA many had already made their own plans to return-to-work and did not wish for the VA to provide them with written documentation of this. Thirdly, there was the potential for recall bias to be introduced when asking participants to recall their work absence over the past months. To examine the potential for the introduction of recall bias a sensitivity analysis on the number of days off work was carried out using the medical record data, which should eliminate recall bias. The findings of this sensitivity analysis again indicated that the number of days off work was reduced in the intervention arm. Lastly, the costs of presenteeism were not included in the economic evaluation because the Stanford Presenteeism Scale used could not be converted into a monetary value. Goetzel *et al* [33] reported that presenteeism accounts for between 18% and 60% of all costs of a range of health conditions. Given that there were significant differences

in measures of presenteeism in favour of the intervention, it is likely that our health economic analyses underestimated the cost-effectiveness of the VA service. In terms of the cost effectiveness of the intervention and the small differences in costs and days off work, there remained some uncertainty around estimates. A larger sample size would be able to reduce this uncertainty, and provide a better cost-effectiveness interpretation. An appropriate threshold for this outcome needs to be determined.

By way of a conservative estimate using data for back and neck pain alone rather than all musculoskeletal pain conditions, 31 million days are lost from work per year in the UK [34]. If a similar brief vocational advice service was implemented widely, it could potentially reduce this figure by 16%, equating to an overall societal cost-saving of about £500 million (216 million days lost per year, amounting to an overall saving of $6 billion for the United States).

Future research

Future research should build upon the intervention provided in the SWAP trial, refining the timing of the intervention to those who have at least 10 days work absence. Given that the results demonstrate benefits in patients with musculoskeletal pain, developing and testing vocational advice services with broader patient groups in primary care such as those with mental health conditions and cardiovascular disease would also be helpful.

Conclusions

SWAP is the first trial to evaluate an intervention embedded in primary care providing early vocational advice, based on biopsychosocial principles, for patients with musculoskeletal pain. The trial demonstrated a reduction in days off work in favour of the vocational advice intervention, an increase in self-efficacy to return-to-work, reduced presenteeism and improved performance at work. Greater economic benefits were seen from the addition of the vocational advice intervention compared to best current primary care alone.

There are no conflicts of interest.

References

1   Confederation of British Industry Fit for Purpose: Absence and Workplace Health
     Survey; 2013 Available from:
     http://www.cbi.org.uk/media/2150120/cbipfizer_absence___workplace_health_2013.
     pdf.

2   Bevan S, Quadrello T, McGee R, Mahdon M, Vavrovsky A, Barham L. Fit For
     Work? Musculoskeletal Disorders in the European Workforce. The Work Foundation;
     2009.

3   Zheltoukhova K. Musculoskeletal Disorders and Work: Results of a survey of
     individuals living with Musculoskeletal Disorders in six European countries; 2013.

4   Waddell G, Burton A.K. Is work good for your health and well-being? The Stationary
     Office; 2006.

5   eumusc.net Musculoskeletal Health in Europe Report V.5. Available from:
     http://www.eumusc.net/myUploadData/files/Musculoskeletal%20Health%20in%20Eu
     rope%20Report%20v5.pdf

6   Black C, Working for a healthier tomorrow. The Stationary Office; 2008.

7   Waddell G, Burton A.K, Kendall N. Vocational Rehabilitation: What Works, for
     Whom, and When? The Stationary Office; 2008

8   Hann M, Sibbald B. General practitioners' attitudes towards patients' health and
     work, 2010-2012, Department for Work and Pensions, Research Report No 835;
     2013.

9   Aust B, Helverskov T, Nielsen M.B.D, Bjorner J.B, Rugulies R, Nielsen K, et al. The Danish national return-to-work program – aims, content, and design of the process and effect evaluation. Scand J Work Environ Health. 2012;38(2):120–133.

10   Nice K, Thornton P. Job Retention and Rehabilitation Pilot: Employers Management of Sickness Absence. Department for Work and Pensions report No 227; 2004.

11   Bishop A, Foster N.E, Thomas E, Hay E.M. How does the self-reported clinical management of patients with low back pain relate to the attitudes and beliefs of health care practitioners? A survey of UK general practitioners and physiotherapists. Pain. 2008; 1-2:187—95.

12   Bishop A, Wynne-Jones G, Lawton S.A, van der Windt D, Main C, Sowden G, et al. Rationale, design and methods of the Study of Work and Pain (SWAP): a cluster randomised controlled trial testing the addition of a vocational advice service to best current primary care for patients with musculoskeletal pain (ISRCTN 52269669), BMC Musculoskelet Disord 2014; 15: 232.

13   Kendall N.A.S, Burton A.K, Main C.J, Watson P.J, on behalf of the Flags Think-Tank: Tackling Musculoskeletal Problems: a Guide for the Clinic and Workplace – Identifying Obstacles Using the Psychosocial Flags Framework. The Stationary Office; 2009.

14   Kendall NA, Linton SJ, Main CJ. Guide to Assessing Psychosocial Yellow Flags in Acute Low Back Pain: Risk Factors for Long-Term Disability and Work Loss. Wellington, New Zealand: Accident Rehabilitation and Compensation Insurance Corporation of New Zealand and the National Health Committee; 1997.

15   Shaw, W.S., van der Windt, D. A, Main, C.J., Loisel, P., Linton, S.J., Decade of the Flags Working Group. 2008. Early patient screening and intervention to address individual level Occupational Factors ("Blue Flags") in Back Disability. J

Occ Rehab. Published online 12/12/08. Springer Science & Business Media. LLC 2008.

16   Dunn K.M, Croft P.R, Classification of low back pain in primary care: Using "Bothersomeness" to Identify the Most Severe Cases. Spine. 2005; 30:1887—1892.

17   Shaw W.S, Reme S.E, Linton S.J, Huang Y.H, Pransky G. 3rd place, PREMUS best paper competition: development of the return-to-work self-efficacy (RTWSE-19) questionnaire--psychometric properties and predictive validity. Scand J Work Environ Health. 2011; 37:109—19.

18   Koopman C, Pelletier K.R, Murray J.F, Sharda C.E, Berger M.L, Turpin R.S. et al. Stanford Presenteeism Scale: Health Status and Employee Productivity. J Occup Environ Med. 2002; 44:14—20.

19   Fournier D.A, Skaug H.J, Ancheta J, Ianelli J, Magnusson A, Maunder M. et al. Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optim Method Soft. 2011; 27: 233—249.

20   Watson P.J, Bowey J, Purcell-Jones G, Gales T. General practitioner sickness absence certification for low back pain is not directly associated with beliefs about back pain. Eur J Pain. 2008; 12:314—320.

21   Hilbe, J. M. sg102: Zero-truncated Poisson and negative binomial regression. Stata Technical Bulletin 47: 37–40. Reprinted in Stata Technical Bulletin Reprints 1999; 8:233—236. College Station, TX: Stata Press, 1999.

22   Weidenhammer W, Linde K, Streng A, Hoppe A, Melchart D. Acupuncture for chronic low back pain in routine care. Clin J Pain. 2007; 23:128—135.

23   Lewis M, Morley S, van der Windt DA, Hay E, Jellema P, Dziedzic K. et al. Measuring practitioner/therapist effects in randomised trials of low back pain and neck pain interventions in primary care settings. Eur J Pain. 2010; 14:1033—1039.

24   Eldridge S.M, Deborah D, Kerry S. Sample size for cluster randomized trials: effect of coefficient of variation of cluster size and analysis method. Int J Epidemiol. 2006; 35:1292—1300.

25   Gomes M, Ng E, Grieve R, Nixon R, Carpenter J, Thomson S. Developing appropriate methods for cost-effectiveness analysis of cluster randomised trials. Medical Decision Making. 2012; 32: 350-361.

26   van Duijn, M, Eijkemans M.J, Koes B.W, Koopmanschap M.A, Burton A.K, Burdorf A. The effects of timing on the cost-effectiveness of interventions for workers on sick leave due to low back pain. Occup Environ Med. 2010; 67:744—750.

27   Burton K, Kendall N, McClusky S, Dibben P. Telephonic support to facilitate return to work: what works, how and when? Department for Work and Pensions. Report No 853, 2013.

28   Palmer K.T, Harris E.C, Linaker C, Barker M, Lawrence W, Cooper C, Coggon D. Effectiveness of community and workplace based interventions to manage musculoskeletal-related sickness absence and job loss: a systematic review. Rheumatology. 2012; 51:230—242.

29   Hillage J, Williams M, Marvell R, Shiels C, Gabbay M, Weston K, Humphreys I. Evaluation of the 2010-13 Fit for Work Service pilots: final report. Department for Work and Pensions, Report No 896, 2015.

30   Rannard A, Gabbay M, Sen D, Riley R, Britt D. Feasibility trial of GP and case-managed support for workplace sickness absence. Prim Health Care Res Dev. 2014; 15:252—261.

31   Sanders T, Wynne-Jones G, Artus M, Ong BN, Foster NE, Acceptability of a vocational advice service for patients consulting in primary care with musculoskeletal pain: A qualitative exploration of the experiences of general practitioners, vocational advisors and patients, *Scandinavian Journal of Public Health,* 2017, *In Press*

32   Coole C, Birks E, Watson P.J, Drummond A. Communicating with Employers: Experiences of Occupational Therapists treating people with musculoskeletal conditions. J Occup Rehabil. 2014; 24:585—595.

33   Goetzel R.Z, Long S.R, Ozminkowski R.J, Hawkins K, Wang S, Lynch W. Health, absence, disability and presenteeism cost estimates of certain physical and mental health conditions affecting U.S employers. J Occup Environ Med. 2004; 46:398—412.

34   Office for National Statistics, Full Report: Sickness Absence in the Labour Market February 2014, http://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/labourproductivity/articles/sicknessabsenceinthelabourmarket/2014-02-25

Figure legends

Figure 1: Model of stepped care provided by the vocational advisor (VA)

Figure 2: CONSORT Flow diagram

Table 1: Baseline characteristics of trial participants by treatment group

	Intervention arm	Control arm
	n=158*	n=180*
Age, mean (SD)	49.5 (9.6)	47.9 (10.7)
Gender, n (%)		
Females	89 (56%)	106 (59%)
Males	69 (44%)	74 (41%)
Duration of symptoms, n (%)		
< 2 weeks	19 (12%)	28 (16%)
2-6 weeks	31 (20%)	49 (28%)
6-12 weeks	28 (18%)	29 (16%)
3-6 months	28 (18%)	31 (18%)
7-12 months	16 (10%)	15 (8%)
> 12 months	35 (22%)	25 (14%)
Time since pain free month, n (%)		
< 3 months	53 (34%)	58 (34%)
4-6 months	12 (8%)	29 (17%)
7-12 months	24 (15%)	22 (13%)
1-3 years	21 (13%)	32 (18%)
> 3 years	46 (30%)	32 (18%)
NRS-Pain average last 2 weeks[a], mean (SD)	6.9 (2.0)	7.0 (1.7)
NRS-Pain least pain last 2 weeks[a], mean (SD)	4.2 (2.4)	3.8 (2.5)
NRS-Pain intensity at present[a], mean (SD)	5.6 (2.5)	5.4 (2.6)
NRS-Pain score summary[a], mean (SD)	5.5 (1.9)	5.4 (1.8)
Bothersomeness, n (%)		

Not at all	0 (0%)	0 (0%)
Slightly	3 (2%)	3 (2%)
Moderately	41 (26%)	42 (23%)
Very much	66 (42%)	85 (47%)
Extremely	48 (30%)	50 (28%)
General health, n (%)		
Excellent	13 (8%)	13 (7%)
Very good	45 (28%)	62 (34%)
Good	61 (39%)	66 (37%)
Fair	31 (20%)	29 (16%)
Poor	8 (5%)	10 (6%)
HADS anxiety[b], mean (SD)	8.0 (4.4)	7.8 (4.1)
HADS depression[c], mean (SD)	6.8 (4.3)	7.0 (4.2)
Working full-time, n (%)	111 (71%)	122 (68%)
Time off work due to pain (past 12 months), n (%)	87 (55%)	113 (63%)
Days off work (past 12 months), mean (range)	15.0 (0-147)	17.8 (0-252)
Has self-certified, n (%)	43 (27%)	57 (32%)
Percent of days off through self-certification,	31%	29%
Has been issued a sick note / fit note, n (%)	60 (38%)	82 (46%)
Percent of days off through sick-certification	69%	71%
Satisfaction with work[d], mean (SD)	6.4 (2.5)	6.4 (2.4)
Performance at work[e], mean (SD)	6.1 (2.6)	6.4 (2.9)
Stanford Presenteeism Scale[f], mean (SD)	18.1 (5.4)	18.0 (5.4)
Self-efficacy – Return to Work[g], mean (SD)	65.9 (27.6)	65.3 (28.8)
Current work situation, n (%)		
Doing usual job	97 (61%)	97 (55%)
On paid annual leave / holiday	3 (2%)	4 (2%)

Working fewer hours	12 (8%)	5 (3%)
Doing lighter duties	7 (4%)	9 (5%)
On paid sick leave	35 (22%)	51 (29%)
On unpaid leave	4 (3%)	11 (6%)
Difficulty managing at work, n (%)		
Not at all	2 (1%)	5 (3%)
Slightly	23 (15%)	34 (19%)
Moderately	52 (34%)	61 (34%)
Very much	45 (29%)	35 (20%)
Extremely	32 (21%)	44 (25%)
NS-SEC		
1	16 (8.9%)	2 (1.2%)
2	35 (19.4%)	40 (26.0%)
3	36 (20.0%)	28 (18.2%)
4	4 (2.2%)	13 (8.4%)
5	9 (5.0%)	12 (7.8%)
6	41 (22.8%)	32 (20.8%)
7	39 (21.7%)	27 (17.5%)
Work is physically demanding	110 (71%)	119 (66%)
Size of organisation >250 staff	44 (29%)	65 (37%)

[a] NRS-Pain scales are 0-10 where 0=no pain, 10=pain as bad as can be;

[b] Pain self-efficacy scale 0-60 where 0=no confidence, 60=highest confidence;

[c] HADS anxiety/depression subscales 0-21 scales where 0=no anxiety/depression, 21=highest anxiety/depression

(clinical cut-offs are given as ≥8 'possible cases' and ≥11 'probable cases');

[d] Satisfaction with work 0-10 NRS scale where 0= not at all satisfied, 10=completely satisfied;

[e] Performance at work 0-10 NRS scale where 0=not at all affected, 10=pain is so bad that unable to do job.

[f] Stanford Presenteeism (6-36 integer scale) where 6=lowest level of presenteeism, 36 highest level of presenteeism;

[g] Self-efficacy Return-to-Work (0-114 scale) where 0=not at all confident, 114=totally confident.

* Not all figures add to the corresponding group totals due to some missing baseline data.

Table 2: Summary of Vocational Advice service delivered

Participants referred to VA service, n (% of intervention group)	120 (76%)
At least one participant contact with the VA	97 (81%)
Number of contact attempts per participant, median (IQR)	4 (2 – 5)
Number of actual contacts per participant, median (IQR)	2 (1 – 3)

Contacts

Total number of participant contact attempts	489
Number of actual participant contacts	226 (37%)
via telephone[a]	202 (89%)
via face-to-face contact[b]	17 (8%)
other (e.g. letter)	7 (3%)
† Duration of telephone call, median (IQR)	13.3 (10 – 20)
‡ Duration of face-to-face visit, median (IQR)	60.0 (35 – 63.5)

Content of vocational advice service

Exploration of health issues	197 (87%)
Exploration of work situation	176 (78%)
Oral information provided	138 (61%)
Assessment of obstacles/'flags' [c]	115 (51%)
Written information provided	20 (9%)
Explored work situation	11 (5%)
Developed Return-To-Work plan	7 (3%)

Number of stakeholder contacts	125

Figures are frequency count (percent) unless otherwise specified.

[a] All 17 face-to-face contacts were with 17 different participants (one of these face-to-face contacts was carried out in the participant's workplace).

[b] Stakeholder contacts were predominantly discharge letters to GPs.

[c] The Flags framework is a system for identifying obstacles to working.

Table 3: Evaluation of the primary outcome measure (days off work) and key secondary outcomes relating to time off work over 4 and 12 months follow up

	4 months				12 months			
	Intervention arm n=119	Control arm n=148	IRR[a] / OR (95% CI)	P-value	Intervention arm n=101	Control arm n=122	IRR[a] / OR (95% CI)	P-value
Days off work[b], mean (SD)	9.29 (21.7)	14.4 (27.7)	0.51 (0.26, 0.99)	0.048	20.3 (40.6)	24.3 (50.7)	0.65 (0.34, 1.25)	0.198
- via self-certification	0.85 (4.11)	0.95 (3.81)	1.14 (0.50, 2.56)	0.759	3.86 (15.0)	1.47 (3.27)	2.97 (1.60, 5.52)	0.001
- via Fit note(s)	8.43 (21.0)	13.5 (27.5)	0.66 (0.46, 0.94)	0.020	16.4 (34.2)	22.9 (50.5)	0.61 (0.41, 0.92)	0.018
Subgroup analysis[c]								
Self-efficacy return-to-work[d]			1.01 (0.87, 1.17)	0.877			1.08 (0.93, 1.26)	0.312
Spinal pain vs. pain in other areas[e]			0.69 (0.27, 1.77)	0.440			0.25 (0.10, 0.62)	0.003
Days-off (prior 12 months)[f]			0.93 (0.78, 1.11)	0.420			0.83 (0.67, 1.03)	0.092
- exceeding 10 days[g]			0.42 (0.17, 1.01)	0.053			0.30 (0.11, 0.83)	0.020

Secondary outcomes

Any reported time off work[h], n (%)	40 (33.6%)	56 (37.8%)	0.64 (0.33, 1.23)	0.182	52 (50.5%)	64 (51.6%)	0.69 (0.34, 1.38)	0.288

Medical record review

Fit note issued[i], n (%)	51 (32.3%)	70 (38.9%)	0.53 (0.25, 1.13)	0.103	52 (50.5%)	64 (51.6%)	0.55 (0.29, 1.04)	0.065
Number of fit notes issued, mean (SD)	0.68 (1.29)	0.94 (1.60)	0.60 (0.35, 1.01)	0.053	1.11 (1.92)	1.51 (2.57)	0.63 (0.37, 1.05)	0.073

[a] Incidence rate ratio (IRR) was the effect of interest (except for self-report time off work (yes/no) and whether a fit note was issued to the participant (yes/no) where the effect of interest was odds ratio (OR)).

[b] Sensitivity analysis of primary outcome (days off work over 4 months, 12 months follow up): (1i) Zero-inflated negative binomial (ZINB) regression with robust variance estimator adjusted for age, gender and practice size [IRR=0.56 (p=0.009) at 4 months, IRR=0.65 (p=0.107) at 12 months]; (1ii) ZINB adjusted for age, gender, practice size plus baseline pain scores and days off over the past 12 months [IRR=0.57 (p=0.004) at 4 months, IRR=0.79 (p=0.391) at 12 months]; (2) Nonparametric (Mann-Whitney U test) comparison of mean ranks (days off work aggregated at cluster (GP-level)) [p=0.343 (4 months), p=0.175 (12 months)]; (3i) Per protocol analysis [IRR=0.52 (p=0.005) at 4 months, IRR=0.55 (p=0.036) at 12 months]; (3ii) Complier average causal effect (CACE) analysis based on two-stage least squares instrumental variable with robust

variance (compliers defined as having at least one contact with the VA (n=97)) [p=0.051 (4 months), p=0.147 (12 months)]. (4) GP practice as random factor (cluster variable) [p=0.019 (4 months), p=0.198 (12 months)].

[c] Subgroup analyses as pre-specified in the published study protocol:

[d] Units denote 10-point increments on the self-efficacy scale;

[e] Days off over 4 months follow up (i) Control group, no spine pain (n=55, mean=10.4, SD 24.7); (ii) Control group, spine pain (n=93, mean=16.8, SD 29.2); (iii) Intervention group, no spine pain (n=42, mean=15.1, SD 26.0); (iv) Intervention group, spine pain (n=77, mean=6.1, SD 18.3); days off over 12 months follow up (i) Control group, no spine pain (n=46, mean=11.8, SD 22.1); (ii) Control group, spine pain (n=76, mean=32.0, SD 60.8); (iii) Intervention group, no spine pain (n=34, mean=32.0, SD 54.0); (iv) Intervention group, spine pain (n=67, mean=14.3, SD 30.6).

[f] Units denote 20-day increments (i.e. about 1 month) on the scale of days off work ([g] additional subgroup analysis requested by TSC).

[h] Time off work (yes/no) – frequency counts (percent) are for participants who reported having had time off work.

[i] Agreement between self-reported time off work (yes/no) and medical record review of issuing of fit note(s) (yes/no) was 70% (187/267) over 4 months and 62% (146/234) over 12 months.

Days off work, QL-QU (90th percentile; max) mean (range) intervention group 4 months: 0-5 (40; 84) via self-certification 0-0 (2; 40) via Fit note 0-0 (40; 84). Intervention group 12 months: 0-15 (80; 210), via self-certification 0-0.3 (8; 126), via Fit note 0-10 (63; 188). Control group 4 months 0-10 (90; 84), via self-certification 0-0 (3; 42), via Fit note 0-10 (63; 188). Control group 12 months 0-3 (75; 252), via self-certification 0-1 (5-19), via Fit note 0-25 (75; 252).

Table 4: Evaluation of secondary outcome measures over 4 and 12 months follow up.

	4 months				12 months			
	Intervention arm	Control arm	MD [a]/OR [b] (95% CI)	P-value	Intervention arm	Control arm	MD [a]/OR [b] (95% CI)	P-value
Pain-related								
NRS-Pain average last 2 weeks, mean (SD)	4.3 (2.8)	5.1 (2.8)	-0.78 [a] (-1.61, 0.04)	0.063	3.6 (3.0)	4.4 (2.4)	-0.76 [a] (-1.82, 0.30)	0.159
NRS-Pain least pain last 2 weeks, mean (SD)	2.9 (2.5)	3.1 (2.5)	-0.20 [a] (-1.05, 0.64)	0.636	2.3 (2.6)	2.4 (2.3)	-0.09 [a] (-1.03, 0.85)	0.854
NRS-Pain intensity at present, mean (SD)	3.3 (2.7)	4.0 (2.9)	-0.63 [a] (-1.58, 0.32)	0.191	2.8 (3.0)	3.6 (2.6)	-0.86 [a] (-1.96, 0.23)	0.122
NRS-Pain score summary, mean (SD)	3.5 (2.5)	4.1 (2.5)	-0.56 [a] (-1.37, 0.24)	0.172	2.9 (2.7)	3.5 (2.3)	-0.59 [a] (-1.56, 0.38)	0.231

Global change, n (%)

Completely recovered	6 (6)	5 (4)	0.87 [b]	0.753	11 (13)	8 (7)	0.96 [b]	0.939
Much improved	18 (18)	33 (27)	(0.35, 2.13)		25 (30)	38 (35)	(0.32, 2.85)	
Somewhat improved	27 (27)	31 (26)			12 (15)	25 (23)		
Same	28 (28)	29 (24)			16 (20)	26 (24)		
Somewhat worse	15 (15)	18 (15)			10 (12)	10 (9)		
Much worse	5 (5)	5 (4)			8 (10)	2 (2)		
Bothersomeness, n (%)			0.82 [b]	0.635			0.44 [b]	0.052
Not at all	2 (2)	3 (2)	(0.36, 1.87)		7 (7)	4 (3)	(0.20, 1.01)	
Slightly	22 (19)	35 (25)			24 (23)	26 (21)		
Moderately	44 (38)	47 (33)			34 (32)	51 (41)		
Very much	30 (26)	39 (27)			27 (26)	38 (30)		
Extremely	17 (15)	18 (13)			13 (12)	6 (5)		

Psychological variables & general health

Pain self-efficacy scale, mean (SD)	41.0 (15.1)	38.0 (14.6)	3.00 [a] (-1.52, 7.53)	0.193	44.7 (14.8)	42.9 (12.2)	1.84 [a] (-3.14, 6.82)	0.470

Illness Perceptions (IPQ-R Short form), n (%)	-	-			-	-		
Identity, median (IQR)	5 (3, 5)	5 (4, 5)	-0.24 [a] (-0.62, 0.14)	0.213	5 (3, 5)	4 (3, 5)	-0.10 [a] (-0.57, 0.38)	0.681
Timeline, n (%)	71 (68.9)	77 (61.1)	0.79 [b] (0.21, 2.97)	0.732	44 (53.0)	66 (60.0)	0.19 [b] (0.04, 0.91)	0.037
Consequences, n (%)	58 (56.3)	64 (50.8)	0.40 [b] (0.09, 1.83)	0.239	34 (40.5)	44 (40.0)	0.36 [b] (0.05, 2.52)	0.304
Personal control, n (%)	49 (48.5)	56 (45.9)	3.41 [b] (1.08, 10.7)	0.036	40 (48.8)	55 (50.5)	1.65 [b] (0.43, 6.32)	0.464
Treatment control, n (%)	70 (69.3)	76 (62.3)	1.27 [b] (0.46, 3.49)	0.639	45 (54.2)	63 (57.3)	0.97 [b] (0.32, 2.90)	0.952
Illness coherence, n (%)	23 (22.8)	29 (23.8)	0.72 [b] (0.20, 2.57)	0.618	12 (14.5)	24 (21.8)	0.11 [b] (0.01, 0.81)	0.031
Timeline cyclical, n (%)	50 (49.5)	60 (49.2)	1.77 [b] (0.58, 5.39)	0.315	30 (36.1)	67 (60.9)	0.17 [b] (0.04, 0.75)	0.019

Emotional representation, n (%)	65 (64.4)	86 (70.5)	0.30 [b] (0.07, 1.22)	0.093	47 (56.6)	59 (54.1)	0.54 [b] (0.11, 2.64)	0.445
HADS anxiety, mean (SD)	6.6 (4.7)	7.9 (4.3)	-1.31 [a] (-2.63, 0.00)	0.050	6.6 (4.1)	7.1 (4.0)	-0.52 [a] (-1.92, 0.87)	0.461
HADS depression, mean (SD)	5.7 (4.2)	6.1 (3.9)	-0.37 [a] (-1.64, 0.91)	0.572	4.7 (3.9)	5.2 (3.8)	-0.47 [a] (-1.81, 0.87)	0.489
General health, n (%)			1.01 [b] (0.33, 3.07)	0.985			0.39 [b] (0.12, 1.26)	0.116
Excellent	8 (8)	8 (6)			4 (5)	4 (4)		
Very good	29 (29)	30 (24)			27 (34)	36 (33)		
Good	32 (32)	54 (43)			29 (36)	46 (42)		
Fair	26 (26)	30 (24)			17 (21)	21 (19)		
Poor	6 (6)	4 (3)			3 (4)	3 (3)		
Work-related								
Stanford Presenteeism Scale, mean (SD)	21.3 (5.4)	19.1 (5.9)	2.23 [a] (0.35, 4.10)	0.020	22.0 (5.6)	20.1 (5.7)	1.89 [a] (-0.24, 4.03)	0.082

Self-efficacy – Return to Work, mean (SD)	81.5 (26.8)	70.1 (27.2)	11.4 [a] (2.97, 19.8)	0.008	82.6 (27.1)	73.7 (24.1)	8.91 [a] (0.04, 17.8)	0.049
Satisfaction with work, mean (SD)	6.4 (2.8)	6.0 (2.3)	0.38 [a] (-0.45, 1.20)	0.369	6.2 (2.6)	6.1 (2.3)	0.06 [a] (-0.83, 0.95)	0.894
Performance at work, mean (SD)	4.1 (2.8)	5.1 (3.0)	-1.05 [a] (-1.96,-0.14)	0.023	3.4 (3.1)	4.6 (2.9)	-1.11 [a] (-2.12,-0.09)	0.032

Descriptive summaries are marginal mean (standard deviation) or frequency count (percent) as appropriate to the type of data being summarised (numerical or categorical, respectively).

[a] MD = Mean Difference (by linear mixed model) / [b] OR = Odds Ratio (by binary/ordinal logit mixed model) adjusted for age, gender and practice size. Attitudes and beliefs (patients) re: work and health will be reported elsewhere to allow the measure to be developed. The content of the GP/NP consultation and questions regarding treatment satisfaction will also be reported separately.

Table 5: Results of the economic evaluation. Values are means (standard deviations) unless stated otherwise

	Intervention arm n = 109	Control arm n = 131
Cost analysis		
Mean (SD) NHS cost (£)	528.34 (1110.49)	480.29 (938.77)
Adjusted mean difference (95% CI) [p-value][b]	48.04 (-209.58 to 305.68) [0.715]	
Mean (SD) Healthcare cost (£)	568.10 (1127.39)	553.32 (976.58)
Adjusted mean difference (95% CI) [p-value][b]	14.78* (-249.76 to 279.33) [0.913]	
Total indirect costs (Benefit) (£)	1636.69 (3671.02)	2257.56 (5233.29)
Adjusted mean difference (95% CI) [p-value][b]	-748** (-2278.45 to 781.44)	
Effectiveness analysis (Work-related outcomes)		
Mean (SD) Days off work	20.26 (40.63)	24.34 (50.67)
Adjusted days off work Mean difference (95% CI's) [p-value][b]	-6.67 (-23.55 to 10.20) [0.438]	
Cost effectiveness and cost-benefit analyses		
ICER NHS perspective	-£7.2 per sick day avoided	
ICER Health care perspective	-£2.2 per sick day avoided	
Net societal benefit	£733 (£748**-£15*)	
Return on Investment (per £1 invested)	£49 (£733/£15*)	

CEA – based on the net monetary benefit (NMB) *; ICER – Incremental cost-effectiveness ratio
.

b Incremental days off work estimated controlling for group and GP Clustering using a GLM regression model, assuming a Gaussian Variance function, an identity Link

Function, and clustered standard errors;

Step 3: Further
face to face
meetings with the
VA

- Targeted advice
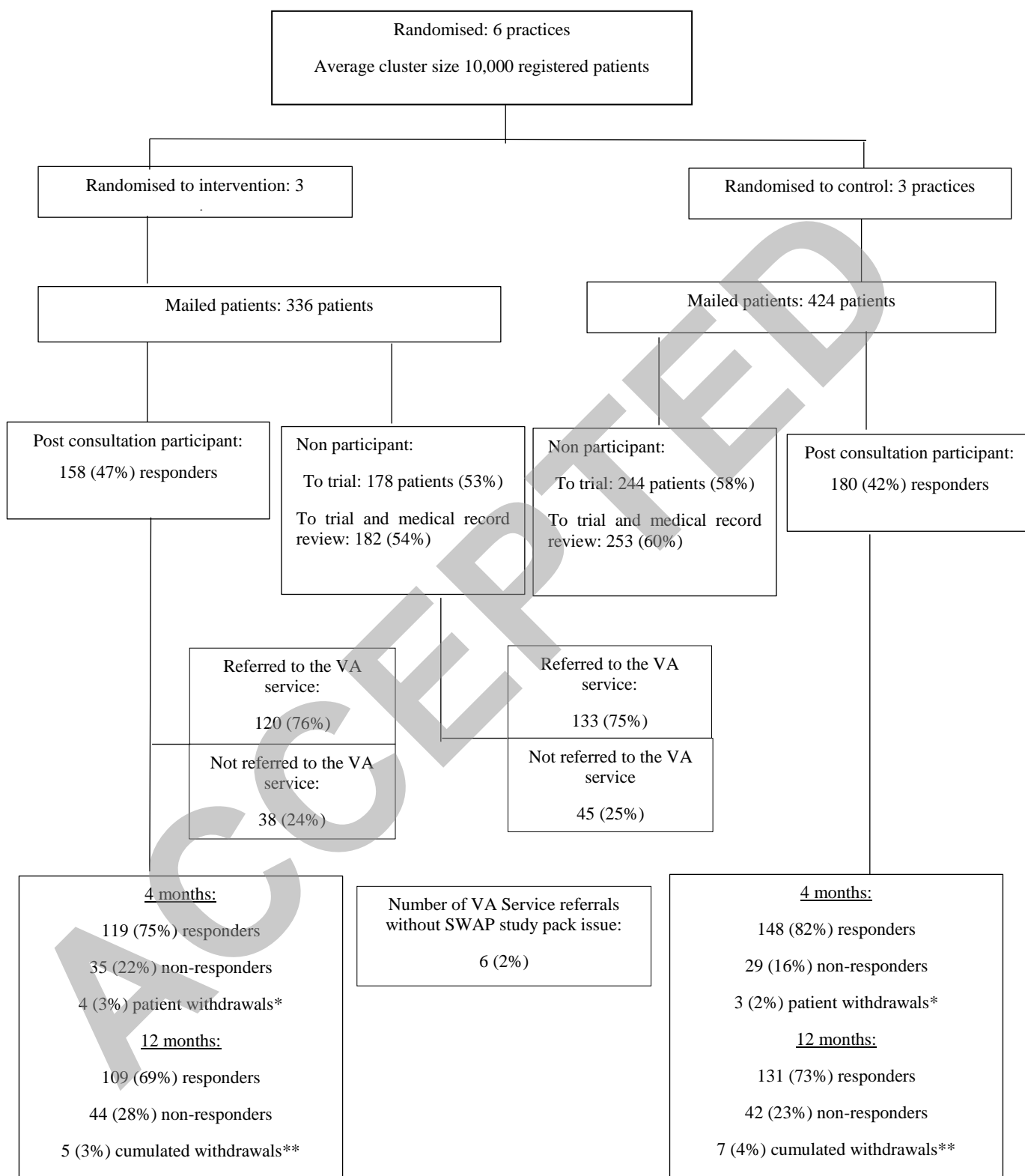- Contact workplace and other services (as required).
- Set new date for RTW

Step 2: Face to
face meeting with
the VA

- Assessment of obstacles to work.
- Develop strategies tackle these.
- Develop return to work plan.

Step 1: Telephone
contact with the
VA

- Initial assessment of beliefs about work and health and obstacles to remaining in/returning to work.
- Discuss date for return to work.

```
                    ┌─────────────────────────────────────┐
                    │      Randomised: 6 practices         │
                    │                                      │
                    │ Average cluster size 10,000 registered patients │
                    └─────────────────────────────────────┘
                                     │
              ┌──────────────────────┴──────────────────────┐
   ┌──────────────────────────────┐         ┌──────────────────────────────┐
   │ Randomised to intervention: 3 │         │ Randomised to control: 3 practices │
   └──────────────────────────────┘         └──────────────────────────────┘
              │                                          │
   ┌──────────────────────────────┐         ┌──────────────────────────────┐
   │  Mailed patients: 336 patients │         │  Mailed patients: 424 patients │
   └──────────────────────────────┘         └──────────────────────────────┘
```

Post consultation participant:

158 (47%) responders

Non participant:

To trial: 178 patients (53%)

To trial and medical record review: 182 (54%)

Non participant:

To trial: 244 patients (58%)

To trial and medical record review: 253 (60%)

Post consultation participant:

180 (42%) responders

Referred to the VA service:

120 (76%)

Not referred to the VA service:

38 (24%)

Referred to the VA service:

133 (75%)

Not referred to the VA service

45 (25%)

4 months:

119 (75%) responders

35 (22%) non-responders

4 (3%) patient withdrawals*

12 months:

109 (69%) responders

44 (28%) non-responders

5 (3%) cumulated withdrawals**

Number of VA Service referrals without SWAP study pack issue:

6 (2%)

4 months:

148 (82%) responders

29 (16%) non-responders

3 (2%) patient withdrawals*

12 months:

131 (73%) responders

42 (23%) non-responders

7 (4%) cumulated withdrawals**

* 7 participant withdrawals between baseline and 4 months follow up: 4 did not wish to take part (2 in the Intervention group and 2 in the Control group); 3 had moved away (2 in the Intervention group and 1 in the Control group).

** 5 participant withdrawals between 4 months and 12 months follow up: 1 did not wish to take part (Intervention group); 2 had moved away (Control group); 1 returned blank questionnaire (Control group); 1 withdrawal reason not known (Control group).