Abstract

Purpose: Exercise (planned, structured, repetitive movement) improves pain and function in people with persistent musculoskeletal pain (PMSK) but adherence is often poor. This systematic review evaluates the evidence from randomised controlled trials (RCTs) on the effectiveness of interventions to improve exercise adherence in people with PMSK and describes the content, context, and theoretical underpinning of behaviour change interventions designed to increase adherence.

Methods: Nine electronic databases were searched from inception dates to August 2017. Studies were included if they were: RCTs which included adults with PMSK ≥three months; ≥one measure of exercise adherence, exercise prescribed to both groups and employed ≥one behaviour change technique (BCT) in the treatment group. Independent data extraction, theory coding, BCT taxonomy coding and quality assessment using Cochrane Risk of Bias (RoB) tool was conducted by two reviewers.

Results: Eight RCTs (five low, three high RoB) met inclusion criteria. Five trials reported between group differences in exercise adherence, favouring the treatment group. Three trials reported theoretical underpinning. There was moderate evidence that five BCTs: social support, goal setting, instruction of behaviour, demonstration of behaviour, practice/rehearsal improved exercise adherence. Interventions employing ≤seven BCTs, unique to those included in the control group, were most effective at enhancing exercise adherence.

Conclusions: Limited moderate quality evidence supports using a small number of BCTs to enhance exercise adherence in people with PMSK. Further research should explore the associations and synergies between BCTs and explicitly report how theory was utilized. This may inform recommendations for healthcare professionals working with this population.
Introduction

Persistent musculoskeletal pain (PMSK) is defined as pain of the axial skeleton that persists longer than expected after onset, usually for at least three months duration (Clinical Standards Advisory Group, 2000). Common diagnoses include low back pain, osteoarthritis and fibromyalgia (Dieppe, 2012). PMSK conditions are challenging to manage and have high personal, health and socioeconomic costs (Reid et al., 2011).

People with PMSK are often referred to a healthcare professional (HCP) and prescribed exercise or physical activity for treatment or management of their pain (World Health Organization, 2018). Exercise is defined as planned, structured and repetitive bodily movement done to improve or maintain physical fitness (Caspersen, Powell, & Christenson, 1985); while physical activity (PA) is any bodily movement that results in energy expenditure (Caspersen et al., 1985). Interventions prescribing exercise or PA improve pain, function and quality of life in individuals with PMSK, including individuals with low back pain (UK Beam, 2004), lower extremity osteoarthritis (Bearne, Walsh, Jessep, & Hurley, 2011) and fibromyalgia (O'Connor et al., 2015). However, the majority of people with PMSK do not experience the benefits of exercise due to poor adherence to their prescribed programme (Peek, Sanson-Fisher, Mackenzie, & Carey, 2016). Treatment adherence has been defined as observing the behaviour endorsed by a HCP for the recommended duration (World Health Organization, 2003) and has been explored in similar health behaviours, such as medication adherence (Osterberg & Blaschke, 2005). However, due to the complexity of health behaviours, the findings from one behavioural domain cannot be easily transferred to another (Bartholomew et al., 2016). Factors influencing adherence may vary depending on the type of behaviour being assessed (Jack, McLean, Klaber, & Gardiner, 2010). It is therefore crucial to be explicit about the behaviour being measured and to distinguish between exercise and PA behaviour. Despite evidence that lifestyle PA, such as walking, improves outcomes (O'Connor...
et al., 2015), current clinical guidelines prioritise the use of exercise programmes for the management of PMSK conditions (Busch et al., 2011; National Institute of Health and Care Excellence, 2016). Therefore, this review focuses on adherence to prescribed exercise.

Evidence suggests that exercise behaviours are modifiable and behaviour initiation and maintenance may be enhanced with behaviour change interventions (Ben-Ami, Chodick, Mirovsky, Pincus, & Shapiro, 2017; Michie, Abraham, Whittington, Mcateer, & Gupta, 2009). However, the factors driving the initiation of a behaviour differ from those contributing to its maintenance (Rothman, 2000). Exercise adherence may encompass both the initiation and maintenance of the behaviour. In this review, the term ‘exercise adherence’ will therefore encompass both of those behaviours.

The identification of behaviour change techniques (BCTs) may inform the development of a behaviour change intervention and recommendations for HCPs working with this population. A taxonomy identified 93 BCTs as the active components in interventions designed to change behaviour (Michie et al., 2013) in order to aid the development, reporting and replication of interventions. In addition, the context within which the BCTs are delivered may also influence effectiveness. For instance, who delivers an intervention, how much training they have received to deliver the intervention, the patient-provider relationship and the environment all should be explored when assessing the effectiveness of behaviour change interventions (Davidson et al., 2003; Drahota et al., 2012). These factors should be reported in the description of trials to aid in intervention replication, and allow comparisons and conclusions to be drawn between interventions (Hoffmann et al., 2014). The Template for Intervention Description and Replication checklist (TIDieR; Hoffmann et al., 2014) has been developed to guide description of interventions.

The TIDieR checklist prompts the reporting of theory to provide rationale for the study protocol. The use of theory is recognized by the Medical Research Council as the first
step in the development of complex interventions (Craig et al., 2013). There is increasing evidence that interventions developed with a theoretical foundation are more effective than interventions without an explicit theoretical underpinning (Glanz, 2015; Prestwich et al., 2014). Nevertheless, there remains conflicting evidence regarding the efficacy of the use of theory to underpin intervention effectiveness and so further research is required (Gourlan et al., 2014; Noar, Benac, & Harris, 2007).

Research into exercise adherence in people with PMSK has been hampered by a lack of standardised outcome measures (Beinart, Goodchild, Weinman, Ayis, & Godfrey, 2013; Geneen et al., 2017). This results in heterogeneous outcome data which limits the ability to compare interventions. Despite this, previous reviews have explored the effectiveness of interventions to improve adherence to prescribed exercise and physiotherapy programmes (Jordan, Holden, Mason, & Foster, 2010; Peek et al., 2016), as well as the use of BCTs in group based self-management programmes (Keogh, Tully, Matthews, & Hurley, 2015). These reviews have suggested that behaviour change interventions may be effective in improving exercise adherence. However, they do not explicitly identify and compare the components of interventions targeting exercise adherence in people with PMSK. This review evaluated the evidence from randomised controlled trials (RCTs) about the effectiveness of the content and context of behaviour change interventions aimed at increasing adherence to prescribed exercise in people with PMSK.

**Objectives**

This systematic review had three objectives: (i) to describe the content (by coding BCTs) and context (informed by the TIDieR framework) used in behaviour change interventions to enhance adherence to prescribed exercise, (ii) to evaluate the effectiveness of BCTs associated with increased adherence and the context in which they were delivered, (iii) to identify the role of theory in these interventions.
Methods

Protocol and registration

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were followed (Moher, Liberati, Tetzlaff, Altman, & the PRISMA group, 2009) and a checklist is available in supplement one. The review protocol is registered on PROSPERO (CRD42016049907) and can be accessed from http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016049907.

Eligibility criteria

RCTs were included in this review if they investigated participants aged 18-65 years. There is an increased risk of comorbidities, particularly serious pathology, which may affect the management of older adults (Greenhalgh, 2006), and the age range selected for this review reflects the range used in global PA guidelines (World Health Organization, 2011). Trials including older or younger participants (ie <18 or >65 years) were included if 80% of participants were between 18-65 years. Participants must have been diagnosed by a HCP with musculoskeletal pain persisting for three months or longer. Musculoskeletal pain was defined as any condition of the axial skeleton (cervical, thoracic or lumbar spine) or any peripheral joints (adapted from Dieppe, 2012). Treatment groups must have received at least one BCT, defined as any effort by the HCP or researchers to change, or support change, of a behaviour. These may include, but were not limited to, goal setting activities or behaviour monitoring. The trial must have included a specific, measurable prescription of exercise (i.e. a set of planned, structured and repetitive movements to be followed for the duration of the intervention). Outcome measures must have included a measure of adherence (defined by the original authors) to the prescribed exercise programme.

Exclusion criteria included (i) trials that only investigated youth and adolescents <18 years or adults aged over 65 years; (ii) people diagnosed with inflammatory conditions (e.g.
rheumatoid arthritis); (iii) control groups including healthy controls, inpatient populations or surgical waiting list patients.

**Information sources**

Electronic databases were searched by one researcher (L.M) for published (CINAHL, PsychINFO, Embase, MEDLINE and the Cochrane Central Register of Controlled Trials) and unpublished (openDOAR, OpenGrey, Web of Science and Ethos) English language trials from their inception dates to 22nd August 2017. The search used combined terms for chronic pain AND behaviour therapy AND exercise and their appropriate MeSH terms (see supplement two for tailored search strategy and all MeSH terms). Search terms were informed by other reviews on chronic/persistent musculoskeletal pain (Hall et al., 2015; Jordan et al., 2010) and health behaviour change (Galea, Weinman, White, & Bearne, 2013; Keogh et al., 2015). The terms PA, exercise and physical fitness are often used interchangeably, therefore all terms were searched in the databases to ensure comprehensiveness. However, regardless of the term used by study authors, only trials meeting the definition for exercise as defined above were included. Secondly, reference lists of included trials were hand searched. Following removal of duplicates, two reviewers (L.M., L.S.) independently screened the titles and abstracts of potentially relevant trials. Full text screening was then conducted by the two reviewers. A third reviewer (E.G.) was available to act as arbiter but was not required.

**Study selection and data extraction**

Data were independently extracted by two reviewers (L.M., L.S.). Information about theoretical constructs and BCTs included in the treatment and control groups were extracted in addition to the standardized information on study design, participant characteristics, and outcomes (such as measure and duration of adherence). Where insufficient information was reported, authors were contacted for further details.
Risk of bias of individual studies and level of evidence

Individual trials were assessed using the Cochrane Risk of Bias Scale (J. Higgins & Green, 2011). The risk of bias (RoB) tool assesses trials across six bias domains (selection, performance, detection, attrition, reporting and any other biases detected) and scores each domain as either high, low, or unclear RoB. However, due to the nature of the included studies, it was difficult to blind participants and the personnel delivering the intervention (Larkin et al., 2015; van Tulder et al., 2001). Therefore, this domain was not included in the final RoB score. In the case of a lack of description of study procedure, it was assumed the task was not carried out and therefore rated as high RoB.

Study quality ratings were determined following thresholds used previously (Bearne, Byrne, Segrave, & White, 2016). RoB was evaluated as high RoB in the presence of three or more sources, or unclear sources of bias, and low RoB if there was evidence for the presence of less than three sources of bias.

Analysis

To assess the content of the interventions, two reviewers (L.M, S.A) coded the BCTs used in the treatment and control groups using the Behaviour Change Taxonomy v1 (Michie et al., 2013). Both reviewers completed the online BCT version one training in recognizing and coding BCTs prior to coding. Past research has shown that control groups often incorporate BCTs and the nature of the control group can influence study findings (Bishop, Fenge-Davies, Kirby, & Geraghty, 2015). Therefore, BCTs were extracted from both the treatment and control groups. In the case of multiple treatment arms, BCTs were assessed separately between arms and compared with the control group. The effectiveness of each BCT (in treatment or control group) in enhancing exercise adherence was calculated using a rating system based on the level of evidence following the recommendations of the Cochrane Back Group (van Tulder, Furlan, Bombardier, & Bouter, 2003):
1. **Strong evidence** Consistent findings in multiple high-quality trials

2. **Moderate evidence** Consistent findings in multiple low-quality trials and/or one high-quality trial

3. **Limited evidence** One low-quality trial

4. **Conflicting evidence** Inconsistent findings in multiple trials

To assess the application of theory in the development and evaluation of the behaviour change interventions, trials were coded using the 19 item Theory Coding Scheme (TCS; Michie & Prestwich, 2010). To test the extent to which theory was used in intervention design, composite scores using the TCS were created (Prestwich et al., 2014; Webb, Joseph, Yardley, & Michie, 2010). Higher scores (with a range from 0-2) indicate stronger theoretical integration in study design. The first score reflects the extent to which each BCT was linked to a theory-relevant construct (items 7-9 in TCS). The second method measures the extent to which the constructs within theory were targeted by BCTs (items 9-11 in TCS). Lastly, an overall theory score was computed, reflecting the use of TCS items that relate to using theory to develop the intervention. This was done by combining the score of the use of theory in developing intervention techniques (TCS 5) with the first two composite scores (Prestwich et al., 2014; Webb et al., 2010).

The contextual aspects of the interventions were extracted following the TIDieR checklist (Hoffmann et al., 2014). This included who delivered the intervention, mode of delivery (e.g. if it was delivered face to face) and where it was delivered.

**Results**

Due to the heterogeneity in intervention design and outcome reporting, conducting a meta-analysis was not possible. The results are summarised qualitatively and presented as a narrative synthesis, following the guidance of Popay and colleagues (2006). As a result of the limited amount of empirical evidence in this field, all trials identified as meeting search criteria are included in the synthesis of results.
Selection process

A total of 1,943 manuscripts were identified from published, peer-reviewed journals. No unpublished studies were identified. Four additional studies were identified through hand searching relevant reference lists. Following the removal of 451 duplicates, 1,492 titles and abstracts were screened and 77 full texts were identified for full screening, of which eight trials met the inclusion criteria (Figure 1).

Study characteristics and participants

Eight randomized controlled trials were included in this review (Coppack, Kristensen, & Karageorghis, 2012; Friedrich, Gittler, Halberstadt, Cermak, & Heiller, 1998; Friedrich, Gittler, Arendasy, & Friedrich, 2005; Harkapaa, Mellin, Jarvikoski, & Hurri, 1990; Harkapaa, Jarvikoski, Mellin, Hurri, & Luoma, 1991; Huyser, Buckelew, Hewett, & Johnson, 1997; Linton, Hellsing, & Bergstrom, 1996; Peterson et al., 2015; Reilly, Lovejoy, Williams, & Roth, 1989; Vong, Cheing, Chan, So, & Chan, 2011) (Table 1). Study sample sizes ranged from 40 to 459 participants, totalling 1,018 participants. All participants in the included trials were aged between 18-65 years. Trials were conducted in: United States (Huyser et al., 1997; Reilly et al., 1989), China (Vong et al., 2011), United Kingdom (Coppack et al., 2012), Sweden (Linton et al., 1996; Peterson et al., 2015), Austria (Friedrich et al., 1998) and Finland (Harkapaa et al., 1990). Six trials included participants with persistent low back pain (Coppack et al., 2012; Friedrich et al., 1998; Harkapaa et al., 1990; Linton et al., 1996; Reilly et al., 1989; Vong et al., 2011), while the remaining two investigated participants with fibromyalgia (Huysper et al., 1997) or whiplash disorders (Peterson et al., 2015). Mean duration of pain symptoms reported ranged from 1.6 years (Peterson et al., 2015) to 14.6 years (Harkapaa et al., 1990). The definition of adherent behaviour differed across trials, and measures of adherence consisted of: diaries/exercise logs (Friedrich et al., 1998; Linton et al.,...
1996; Vong et al., 2011), questionnaires (Coppack et al., 2012; Huyser et al., 1997), or number of exercise sessions completed (Harkapaa et al., 1990; Reilly et al., 1989). Length of follow up ranged from 15 days (Coppack et al., 2012) to five years (Friedrich et al., 2005).

[Insert Table 1 Here]

**Risk of bias**

The summary of the RoB assessment can be found in Table 2. The two reviewers (L.M., L.S.) had good agreement and any discrepancies were resolved following discussion. Five studies were assessed as low RoB (Coppack et al., 2012; Friedrich et al., 1998; Huyser et al., 1997; Peterson et al., 2015; Vong et al., 2011), and three as high RoB (Harkapaa et al., 1990; Linton et al., 1996; Reilly et al., 1989).

[Insert Table 2 Here]

**Content of the interventions**

Five of the eight included trials reported greater between group adherence favouring the treatment groups (Coppack et al., 2012; Harkapaa et al., 1990; Linton et al., 1996; Reilly et al., 1989; Vong et al., 2011). The standardized mean difference was calculated and presented where possible in Table 1 (forest plot in supplement three). Three of the five trials reporting significant between group differences were considered to be high RoB (Harkapaa et al., 1990; Linton et al., 1996; Reilly et al., 1989), while the non-significant differences were all found in low RoB trials (Friedrich et al., 1998; Huyser et al., 1997; Peterson et al., 2015).

There was good BCT coding agreement between reviewers for treatment groups (75.3%) and control groups (71.6%). Overall, there were 30 different BCTs (out of a possible 93) identified in the eight trials (Table 3). The number of BCTs varied from three to eleven (median 7) in the treatment group and one to six (median 3.5) in the control group. **Instruction on how to perform behaviour, demonstration of behaviour and behaviour practice/rehearsal** were the most commonly employed BCTs. In the treatment group, these three BCTs were
observed in 6/8, 5/8 and 5/8 trials respectively. In the control group, they were observed in 7/8, 4/8, and 5/8 trials respectively.

Based on the Cochrane Back Group rating system (van Tulder et al., 2003) none of the BCTs produced strong evidence for their effectiveness in enhancing exercise adherence. There was evidence that social support (unspecified), goal setting (behaviour), instruction of behaviour, demonstration of behaviour and behaviour practice/rehearsal were moderately effective at enhancing exercise adherence (see article by Michie et al., 2013 for full descriptions of these BCTs).

Two low RoB trials (Friedrich et al., 2005; Peterson et al., 2015) reporting non-significant between group difference in exercise adherence employed the most BCTs within the treatment groups (9-11 BCTs). Additionally, both the treatment and control groups in these two trials received three of the BCTs that had the greatest evidence of effectiveness to enhance exercise adherence (instruction of behaviour, demonstration of behaviour, and behaviour practice/rehearsal) potentially confounding intervention effectiveness.

Three high RoB trials measured adherence to exercise for six months or more. Reilly et al (1989) and Linton et al (1996) reported greater between group exercise adherence favouring the treatment group after six months and Harkapaa et al (1991) observed greater between group exercise adherence after 1.5 years. These trials employed seven or less BCTs in the treatment group and no more than two BCTs within the control group.

The content of the prescribed exercise programmes varied. All were tailored to the participant and delivered face to face by a HCP. Three (two low, one high RoB) trials included varied exercise progressions based on either participant progress or preference (Coppack et al., 2012; Reilly et al., 1989; Vong et al., 2011) and reported significant between group adherence in the treatment group. Two trials (one low, one high RoB) involved group
training (Coppack et al., 2012; Harkapaa et al., 1990), and both reported greater adherence in
treatment groups compared to control.

[Insert Table 3 here]

Context of the interventions

Three trials (Friedrich et al., 1998; Peterson et al., 2015; Vong et al., 2011), administered
interventions solely by a physiotherapist/physical therapist, one trial (Huyser et al., 1997) by a
physician, one trial (Linton et al., 1996) by a psychologist, one trial (Harkapaa et al., 1990) by
a combination of all three, and two trials (Coppack et al., 2012; Reilly et al., 1989) were
delivered by exercise specialists. No included trials provided specific detail regarding the
training offered or completed by the HCPs.

Five low RoB trials delivered the behaviour change programme alongside the
prescribed exercise programme in a clinic. Two trials reported greater adherence in the
treatment groups (Coppack et al., 2012; Vong et al., 2011), while three trials (Friedrich et al.,
1998; Huyser et al., 1997; Peterson et al., 2015) did not report a significant between group
difference. Three trials (high RoB) delivered the behaviour change and exercise programme in
a community health centre and reported greater between group adherence favouring the
treatment group (Harkapaa et al., 1990; Linton et al., 1996; Reilly et al., 1989).

The five trials (two low, three high RoB) reporting significant between group
difference delivered the interventions between two to seven days per week (Coppack et al.,
2012; Harkapaa et al., 1990; Linton et al., 1996; Reilly et al., 1989; Vong et al., 2011). The
duration of the sessions varied between 30 minutes (Coppack et al., 2012; Vong et al., 2011)
to over two hours (Linton et al., 1996) (Table 3).

Theoretical basis of included interventions

Of the eight trials included, only three trials (two low, one high RoB) reported using theory to
guide the intervention. Coppack et al (2012) scored the maximum TCS summary score of five
utilizing tenets of Personal Construct Theory (Kelly, 1963). Constructs of Health Locus of Control (Lefcourt, 1981) were reported by Harkapaa et al (1990), which scored a TCS summary score of three, and a combination of the Social Cognitive Theory (Bandura, 1986) and the Transtheoretical Model (Glanz, Rimer, Viswanath, & Orleans, 2008) informed the development of the intervention by Peterson et al. (2015) and scored a TCS summary score of two (see supplement four for full coding).

**Discussion**

To the authors’ knowledge, this is the first systematic review to evaluate the content and context of behaviour change interventions administered to improve exercise adherence in people with PMSK. There is moderate quality evidence from two trials with low RoB (Coppack et al., 2012; Vong et al., 2011) and three with high RoB (Harkapaa et al., 1991; Linton et al., 1996; Reilly et al., 1989) that the inclusion of a behaviour change intervention incorporating BCTs enhances adherence to prescribed exercise compared to a control group. Furthermore, there is evidence to suggest that the type, amount, and delivery of BCTs will influence the level of adherence. Five BCTs were found to have a moderate level of evidence to support adherence including:

- Social support (unspecified)
- Goal setting (behaviour)
- Instruction of behaviour
- Demonstration of behaviour
- Behaviour practice/rehearsal

Trials implementing these BCTs within their treatment groups produced significant between group differences in favour of the treatment groups. Furthermore, the presence of these BCTs in the control group were associated with higher adherence rates in the control groups in three trials (Friedrich et al., 2005; Huyser et al., 1997; Peterson et al., 2015).
There is some evidence to suggest that trials employing seven or less BCTs have greater effectiveness than those employing more; the two trials (low RoB) employing the most BCTs to treatment groups found no significant between group difference in exercise adherence (Friedrich et al., 2005; Peterson et al., 2015). There is conflicting evidence within the literature regarding the optimum number of BCTs. Our findings concur with the findings of Michie et al (2009), who report that administering more BCTs did not result in higher PA levels. Conversely Bishop (2015) reported greater adherence in interventions with a higher number of BCTs. However, trials with control groups containing a low number of BCTs compared to the treatment group were associated with better outcomes, particularly if the control group contained unique BCTs from the treatment group (Bishop et al., 2015).

Similarly, in our review the three trials (all assessed as high RoB) reporting the highest exercise adherence after six months had less than three BCTs within their control group (Harkapaa et al., 1991; Linton et al., 1996; Reilly et al., 1989). This suggests that the disparity between the number of BCTs used in the treatment and control groups may influence results. The possible cumulative or confounding effects of these BCTs should be investigated to further explore their influence (Michie et al., 2009). Furthermore, this may suggest that accurately and effectively delivering a small number of BCTs may be more important than administering numerous BCTs poorly. The use of BCTs to enhance adherence to treatment has increased (Spetch & Kolt, 2001), however the HCPs providing these treatments may not obtain formal training in BCT delivery (Arvinen-Barrow, Penny, Hemmings, & Corr, 2010), which is likely to be linked to their ability to deliver them well. There was a lack of detail reported regarding the amount of training provided for treatment delivery, or of any assessment of fidelity. This may call into question whether the behaviour change programmes were administered consistently or whether it is feasible for the healthcare professional to provide this type of treatment within their practice.
Our review found that many trials employed the same BCTs in both the treatment and control groups and reported no significant differences between the two groups. This suggests that study design may have influenced results, since monitoring BCTs delivered in usual care or in waiting list control groups can be challenging. This should be considered when designing future studies so that BCTs within control groups can be accurately recorded. The five trials reporting significant between group differences in exercise adherence investigated populations with persistent low back pain, however, there is no evidence to suggest that pain duration was related to adherence. HCPs, such as physiotherapists, should be aware of the value of including BCTs within practice, to initiate or maintain exercise adherence in people with PMSK, regardless of the duration of symptoms.

The type of exercise programme prescribed to participants varied across the trials. The trials producing a significant between group difference all implemented a progressive, graded exercise programme to the treatment group. Three trials reporting significant between group differences administered the programme in a community health club (Harkapaa et al., 1991; Linton et al., 1996; Reilly et al., 1989), and two trials produced significant between group differences with group exercise training (Coppack et al., 2012; Harkapaa et al., 1991). Based on the Hedonic Principle, people will maximise what feels good, and be driven by enjoyment and positive affect (E. T. Higgins, 2012). Enjoyment has been found to be a strong mediator of exercise adherence in people with musculoskeletal pain (Hagberg, Lindahl, Nyberg, & Hellénius, 2009). This review did not include trials that measured enjoyment but this provides some evidence that patients with pain conditions may have stronger adherence rates if they experience more tailored and varied exercise programmes. Additionally, given the effectiveness of group-based physiotherapy programmes on pain outcomes (O'Keeffe, Hayes, Mccreesh, Purtill, & O'Sullivan, 2016), further research should explore the integration of the BCTs within group-based classes.
Participants in the treatment group had more behaviours to adhere to than those in the control groups, as they were asked to participate in both the behaviour change and the prescribed exercise programme. This could have been burdensome for the participants and the personnel delivering the intervention. There was heterogeneity in who administered the treatment, with minimal evidence to support one provider over the other. A viable programme that can be sustained by both patients and HCPs is required to enhance adherence in people with PMSK. Behaviour change interventions may benefit from paying attention to the mode of delivery and the context in which the content is delivered.

There is increasing evidence suggesting that interventions with a clear theoretical underpinning are more effective than those without, particularly in the health behaviour change literature (Glanz & Bishop, 2010). However, conflicting evidence persists within the literature and the current review. Only one intervention in this review explicitly used theory to inform both design and delivery. Furthermore, some of the included trials evaluated interventions which were not explicitly underpinned by theory and reported greater between group differences in exercise adherence favouring the treatment group. Due to these findings and the discrepancies within the literature, questions remain about whether a theoretical underpinning is associated with intervention effectiveness. There needs to be more accurate reporting of the use of theory to support interventions in order to assess whether it is a prerequisite for success.

**Methodological considerations**

There are limitations to this review that should be noted. Due to the small number of included trials and the fact that three of these had a high RoB, findings are tentative and must be treated with caution. It should be noted that the current findings only provide an indication of possible associations between the BCTs and intervention effectiveness and potential dose response relationships. None of the trials included in this review reported conducting an
intervention fidelity assessment, and so it is not known if the interventions, including the BCTs, were delivered consistently and as planned. This may have led to the inaccurate reporting and interpretation of results. Treatment fidelity assessments allow researchers to detect any deviations or errors in the trial design and implementation. Higher levels of treatment fidelity have been associated with better trial retention rates and treatment outcomes (Borrelli, 2011).

Within this review there is also the risk of inaccurate coding of the BCTs, especially when poor intervention reporting, or description, make some techniques difficult to classify. While this effect was minimised in the current review by ensuring consistent interpretation between the two coders, it highlights the need for better reporting in behaviour change literature. The National Institute for Health and Clinical Excellence (2007) guidelines recommend interventions are focused on those behaviours that need modifying. The measure of exercise adherence was often incidental to the overall study objectives, which may have affected the reporting of the intervention and impact on adherence levels.

The methods of reporting exercise adherence as an outcome varied across included trials, limiting comparisons and eliminating the option of conducting a meta-analysis. The included studies all used self-reported measures of exercise adherence that lacked standardisation. This may have led to inaccurate reporting of exercise levels due to recall or self-report bias (Prince et al., 2008). There is a lack of valid and reliable measures of exercise adherence (Beinart et al., 2013) and only a few have undergone psychometric testing (see Newman-Beinart et al., 2017). This highlights the need for standardized, validated measures to move the research forward.

Additionally, this review only included interventions with an exercise prescription, and not general PA. PA, such as walking (O'Connor et al., 2015) and habitual PA (Ben-Ami
et al., 2017), have been found to decrease pain in this population. Future reviews should further explore adherence to PA interventions in a population with PMSK.

**Conclusion**

This review found that there was moderate evidence that five BCTs: social support, goal setting, instruction of behaviour, demonstration of behaviour, and behavioural practice/rehearsal improved exercise adherence in people with PMSK when compared to a control group. Treatment groups including seven or less BCTs, which were unique to any BCTs used in the control group, were most effective at enhancing exercise adherence. HCPs should consider incorporating BCTs into the prescription of exercise for people with PMSK and observed patterns could be used for hypothesis testing in future intervention development. This may aid progress in the field by gathering evidence informing what is delivered, by whom, and where it is best implemented.

**Conflicts**

No potential conflict of interest reported.
References


example of trials to increase adherence to physical activity in musculoskeletal pain.

*Psychology & Health, 30*(1), 104-121.


Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., . . . Wood, E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior...


Figure 1: PRISMA diagram

- Records identified through database searching (n = 1943):
  - Embase n=1129
  - CINAHL n=126
  - Medline n=431
  - PsychINFO n=93
  - Cochrane n=164

- Additional records identified through other sources (n = 4)

- Records after duplicates removed (n = 1492)

- Records screened (n = 1492)
  - Full-text articles assessed for eligibility (n = 77)
    - Full-text articles excluded, with reasons (n = 69)
      - Population=18
      - Intervention=48
      - Outcome=3
  - Studies included in qualitative synthesis (n = 8)
  - Studies included in quantitative synthesis (meta-analysis) (n = 0)
## Table 1: Study characteristics

<table>
<thead>
<tr>
<th>First author</th>
<th>Diagnosis</th>
<th>Pain duration (years)*</th>
<th>N (% Female)</th>
<th>Age *</th>
<th>Measure of Adherence</th>
<th>Length of follow up</th>
<th>Results *</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Treatment</td>
<td>Control</td>
<td></td>
<td>How adherence is calculated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huyser</td>
<td>Fibromyalgia</td>
<td>11.7 (9)</td>
<td>13.2 (9.4)</td>
<td>54 (not reported)</td>
<td>Participant self-report (questionnaire)</td>
<td>6 weeks</td>
<td>4 (1.5)</td>
</tr>
<tr>
<td>(1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vong</td>
<td>Low back pain</td>
<td>3.5 (4.7)</td>
<td>4.3 (6)</td>
<td>76 (63)</td>
<td>Participant self-report (diary)</td>
<td>1 month</td>
<td>12.9 (7.2)</td>
</tr>
<tr>
<td>(2011)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coppack</td>
<td>Low back pain</td>
<td>2.6 (0.3)</td>
<td>32 (6)</td>
<td>33 (7.9)</td>
<td>Sports Injury Rehab Scale (SIRAS)</td>
<td>15 days</td>
<td>13.7 (1.6)</td>
</tr>
<tr>
<td>(2012)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Max 15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peterson</td>
<td>Whiplash associated disorder</td>
<td>1.6 (0.7)</td>
<td>1.7 (0.7)</td>
<td>216 (66)</td>
<td>Physiotherapist and participant self-report (diary)</td>
<td>6 months</td>
<td>53%</td>
</tr>
<tr>
<td>(2015)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>% participants completing &gt;50% prescribed home exercise sessions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friedrich</td>
<td>Low back pain</td>
<td>4.2 (4.1)</td>
<td>3.8 (3.7)</td>
<td>93 (58)</td>
<td>Staff and participant self-report (diary)</td>
<td>5 years</td>
<td>3.5 (2)</td>
</tr>
<tr>
<td>(1998,2005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Years participating in prescribed exercises 3x/week</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>Condition</td>
<td>Exercise Type</td>
<td>M (SD)</td>
<td>N (SD)</td>
<td>Outcome Measure</td>
<td>Time Period</td>
<td>% Adhering</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------</td>
<td>---------------</td>
<td>----------------</td>
<td>----------------</td>
<td>-------------------------------------</td>
<td>-------------</td>
<td>------------</td>
</tr>
<tr>
<td>Harkapaa</td>
<td>Low back pain</td>
<td>Physician self-report (diary)</td>
<td>14.6 (not reported)</td>
<td>459 (37)</td>
<td>% participants performing faultless exercise</td>
<td>1.5 years</td>
<td>51%</td>
</tr>
<tr>
<td>(1990,1991)</td>
<td></td>
<td></td>
<td>13.4 (not reported)</td>
<td>45 (not reported)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reilly</td>
<td>Low back pain</td>
<td>Gym staff report (diary)</td>
<td>Not reported</td>
<td>40 (50)</td>
<td>Number of sessions attended during study</td>
<td>6 months</td>
<td>90.8(3.3)</td>
</tr>
<tr>
<td>(1989)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linton</td>
<td>Low back pain</td>
<td>Participant self-report (diary)</td>
<td>Not reported</td>
<td>48 (42)</td>
<td>% participants completing exercise 2x/week</td>
<td>6 months</td>
<td>52%</td>
</tr>
<tr>
<td>(1996)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Presented as M (SD) unless otherwise indicated
b Total 95% CI 1.12 [-0.15, 2.39] Heterogeneity: Chi²=90.54, df=5 (P< 0.00001)
* p >.05 between group difference
### Table 2: Risk of bias of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Random sequence generation (selection bias)</th>
<th>Allocation concealment (selection bias)</th>
<th>Blindness of outcome assessment (detection bias)</th>
<th>Incomplete outcome data (attrition bias)</th>
<th>Selective reporting (reporting bias)</th>
<th>Summary Score *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coppack</td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /></td>
<td>Low</td>
</tr>
<tr>
<td>Friedrich</td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /></td>
<td>Low</td>
</tr>
<tr>
<td>Hakapaa</td>
<td><img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /></td>
<td>High</td>
</tr>
<tr>
<td>Huyser</td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /></td>
<td>Low</td>
</tr>
<tr>
<td>Lirton</td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /></td>
<td><img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /></td>
<td><img src="image" alt="High risk of bias" /></td>
<td>High</td>
</tr>
<tr>
<td>Peterson</td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /></td>
<td>Low</td>
</tr>
<tr>
<td>Reilly</td>
<td><img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /></td>
<td><img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /> <img src="image" alt="High risk of bias" /></td>
<td><img src="image" alt="High risk of bias" /></td>
<td>High</td>
</tr>
<tr>
<td>Vong</td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /> <img src="image" alt="Low risk of bias" /></td>
<td><img src="image" alt="Low risk of bias" /></td>
<td>Low</td>
</tr>
</tbody>
</table>

*Summary bias score calculated as high where ≥3 sources or unclear and low where <3 RoB

- ![Low risk of bias](image) Low risk of bias
- ![High risk of bias](image) High risk of bias
- ![Unknown risk of bias](image) Unknown risk of bias
Table 3: Intervention components

<table>
<thead>
<tr>
<th>First author</th>
<th>Intervention</th>
<th>Behaviour change programme</th>
<th>Exercise programme</th>
<th>Frequency of in-person sessions</th>
<th>Duration of in-person sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Huyser</td>
<td>Biofeedback/relaxation</td>
<td>(1.6), (2.6), (4.1), (6.1), (8.1), (13.2)</td>
<td>(4.1), (6.1), (8.1)</td>
<td>Physician</td>
<td>Clinic</td>
</tr>
<tr>
<td>Vong</td>
<td>Motivation enhancement</td>
<td>(3.1), (9.1), (9.3)</td>
<td>(1.4), (2.2), (4.1), (8.1)</td>
<td>Physical therapist</td>
<td>Clinic</td>
</tr>
<tr>
<td>Coppack</td>
<td>Goal setting therapy</td>
<td>(1.1), (1.3), (1.5), (1.6), (1.7), (4.1), (6.1), (8.1)</td>
<td>(2.1), (4.1), (6.1), (8.1)</td>
<td>Exercise specialist</td>
<td>Clinic</td>
</tr>
<tr>
<td>Peterson</td>
<td>Neck specific exercise and behaviour enhancement therapy</td>
<td>(1.1), (1.4), (4.1), (5.4), (6.1), (8.1), (8.6), (8.7), (12.6)</td>
<td>(4.1), (5.1), (6.1), (8.1), (8.6), (8.7)</td>
<td>Physio</td>
<td>Clinic</td>
</tr>
<tr>
<td>Study Source</td>
<td>Intervention Description</td>
<td>Motivation</td>
<td>Physiotherapy</td>
<td>Setting</td>
<td>Counselling</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------</td>
<td>------------</td>
<td>---------------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td><strong>Friedrich (1998, 2005)</strong>&lt;br&gt;<strong>Motivation enhancement and exercise programme</strong>&lt;br&gt;(1.2), (1.4), (1.8), (2.3), (4.1), (5.1), (7.1), (8.7), (10.3), (10.4), (10.11)</td>
<td>(4.1), (6.1) (8.1), (8.6), (8.7)</td>
<td>Physio</td>
<td>Clinic</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Harkapaa (1990, 1991)</strong>&lt;br&gt;<strong>Back school with relaxation, and counselling</strong>&lt;br&gt;(2.2), (3.1), (4.1), (6.1), (8.1), (12.6)</td>
<td>(2.2), (4.1)</td>
<td>Psychologist and physician</td>
<td>Local health club</td>
<td>Yes</td>
<td>Not clear</td>
</tr>
<tr>
<td><strong>Reilly (1989)</strong>&lt;br&gt;<strong>Personalised and monitored exercise programme</strong>&lt;br&gt;(2.2), (4.1), (6.1), (8.1)</td>
<td>(4.1)</td>
<td>Exercise specialist</td>
<td>Local health club</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Linton (1996)</strong>&lt;br&gt;<strong>Exercise and intensive counselling</strong>&lt;br&gt;(1.1), (1.2), (1.4), (1.5), (1.9), (7.1), (8.7)</td>
<td>(1.4), (12.5)</td>
<td>Behavioural psychologist</td>
<td>Not reported</td>
<td>Yes</td>
<td>Not clear</td>
</tr>
</tbody>
</table>
(1.1) Goal setting (behaviour), (1.2) Problem solving, (1.3) Goal setting (outcome), (1.4) Action planning, (1.5) Review behaviour goals, (1.6) Discrepancy between current behaviour and goal, (1.7) Review outcome goal(s), (1.8) Behavioural contract, (1.9) Commitment, (2.1) Monitoring of behaviour without feedback, (2.2) Feedback on behaviour, (2.3) Self-monitoring of behaviour, (2.6) Biofeedback, (3.1) Social support (unspecified), (4.1) Instruction on how to perform a behaviour, (5.1) Information about health consequences, (5.4) Monitoring of emotional consequences, (6.1) Demonstration of behaviour, (7.1) Prompts/cues, (8.1) Behavioural practice/ rehearsal, (8.6) Generalisation of a target behaviour, (8.7) Graded tasks, (9.1) Credible source, (9.3) Comparative imagining of future outcomes, (10.3) Non-specific rewards, (10.4) Social reward, (10.11) Future punishment, (12.5) Adding objects to the environment, (12.6) Body changes, (13.2) Framing/reframing

Note: Bolded codes signify BCTs used in both treatment and control conditions
*Trials observing significant in-between differences in control and treatment groups