Feasibility and preliminary effectiveness on pain and self-efficacy of a multidisciplinary care programme for generalised osteoarthritis: a concurrent randomised multiple-baseline single-case study.

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2012-001161
Article Type:	Research
Date Submitted by the Author:	14-Mar-2012
Complete List of Authors:	Hoogeboom, Thomas; Sint Maartenskliniek, Department of Rheumatology; Maastricht University Medical Centre, Department of Epidemiology Kwakkenbos, Linda; Sint Maartenskliniek, Department of Rheumatology Rietveld, Leonie; Maartenskliniek, Department of Rheumatology den Broeder, Alfons; Sint Maartenskliniek, Department of Rheumatology de Bie, Rob; Maastricht University Medical Centre, Department of Epidemiology van den Ende, Cornelia; Sint Maartenskliniek, Department of Rheumatology
Primary Subject Heading:	Rehabilitation medicine
Secondary Subject Heading:	Rehabilitation medicine, Evidence based practice
Keywords:	Rehabilitation medicine < INTERNAL MEDICINE, Rheumatology < INTERNAL MEDICINE, Musculoskeletal disorders < ORTHOPAEDIC & TRAUMA SURGERY

Feasibility and preliminary effectiveness on pain and self-efficacy of a multidisciplinary care programme for generalised osteoarthritis: a concurrent randomised multiple-baseline single-case study.

Short title: Non-pharmacological care in generalized osteoarthritis

Hoogeboom TJ, Kwakkenbos L, Rietveld L., den Broeder AA, de Bie RA, van den Ende CHM.

[§]*Corresponding author:*

Tel: +31 24 327 2718,

Fax: +31 24 365 9154.

E-mail: t.hoogeboom@maartenskliniek.nl

Address for reprints requests:

Sint Maartenskliniek, ReumaResearch

PO Box 9011, 6500 GM, Nijmegen, the Netherlands.

Grant supporter: The study was financed by the Sint Maartenskliniek Nijmegen and Woerden, the Netherlands.

Competing interests: Authors have no competing interests to declare.

Manuscript word count: 2999 words.

Abstract

Objectives. To evaluate the feasibility and to preliminarily evaluate the effectiveness of a 12-week multidisciplinary non-pharmacological intervention in patients with generalised osteoarthritis (GOA).

Design. A randomised, concurrent, multiple-baseline, single-case design. During the baseline period, the intervention period, and the post-intervention period, all participants completed several health outcomes twice a week on visual analogue scales.

Setting. Rheumatology, outpatient department of a specialized hospital in the Netherlands.

Participants. One man and four women (age 51 to 76) diagnosed with GOA.

Primary outcome measures. To assess feasibility we assessed the number of drop-outs and adverse events, adherence rates, and patient satisfaction.

Secondary outcome measures. To assess effectiveness preliminarily we assessed pain and self-efficacy. Effectiveness was preliminarily assessed using visual data inspection and randomisation tests.

Results. The intervention was feasible in terms of adverse events (none) and adherence rate, but not in terms of participant satisfaction with the intervention. Visual inspection of the data and randomisation testing demonstrated no effects on pain (p = 0.93) or self-efficacy (p = 0.85).

Conclusions. The results of the present study indicate that the proposed intervention for patients with GOA was insufficiently feasible and effective. The data obtained through this multiple-baseline study has highlighted several areas in which the therapy programme can be optimised.

Article summary

Article focus:

- To evaluate the feasibility the effectiveness of a 12-week multidisciplinary non-pharmacological intervention in patients with generalised osteoarthritis (GOA).
- To preliminarily evaluate the effectiveness of a 12-week multidisciplinary non-pharmacological intervention in patients with GOA.

Key messages:

- To date no studies are available that evaluate non-pharmacological care in individuals with GOA.
- The intervention evaluated in the present study appeared both insufficiently feasible and effective for patients with GOA.
- Several areas in which the therapy programme could be optimised were highlighted.

Strengths and limitations of this study:

- A multiple-baseline single-case design is particularly successful in demonstrating immediate effects, whereas we studied changes in health behaviour.
- Inherent to the design of the study is lower external validity due to the small number of included participants.

Funding statement:

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors

Competing interests:

All authors declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work.

Author's contribution:

Substantial contribution to the conception and design of the study: TJH, LK, LR, AAB, RAB and CHME.

Substantial contribution to the acquisition of the data: TJH, LR

Substantial contribution to the analysis and interpretation of the data: TJH, LR, LK, CHME, AAB and CHME.

Provided intellectual content whilst drafting the article: TJH, LR, LK, CHME, AAB and CHME.

Approved the final version to be published: TJH, LR, LK, CHME, AAB and CHME.

Data sharing statement:

No additional data available.

Introduction

A growing body of evidence shows that individuals with established osteoarthritis who also report joint-pain comorbidities - often referred to as generalised osteoarthritis (GOA) - represent a relatively large subgroup of patients [1-4]. It has been suggested that these people might be in need of more intensive treatment options than patients with single joint complaints [1,5]. To the best of our knowledge, however, there are no studies that evaluate non-pharmacological care in individuals with GOA [5], warranting the development and evaluation of such a treatment programme. Therefore, we conceptualised a non-pharmacological treatment programme following a previously-described systematic procedure [6]. The intervention was based on recommendations for the management of hip and knee osteoarthritis [7-9], and was tailored to the needs of patients with joint-pain comorbidities [1].

Before evaluating such an intervention in a randomised clinical trial, a pilot study is recommended [10], since evaluations are often undermined by problems of acceptability, compliance, delivery of the intervention, recruitment and retention, and smaller-than-expected effect sizes [11]. A useful study design for pilot interventions is the multiple-baseline single-case design, as it allows researchers to test the feasibility of the intervention and to make a preliminary assessment of its effectiveness with a low number of participants [12]. In a multiple-baseline design, the intervention is introduced to subjects after randomly-assigned baseline periods of different lengths, and an effect is demonstrated if the measured outcome only changes after the intervention has been introduced [13].

The primary aim of our study was to evaluate the feasibility of a complex multi-disciplinary intervention in patients with GOA. Our secondary aim was to preliminarily assess the effectiveness of this intervention on pain and self-efficacy.

Methods

Participants

Men and women, 40 years or older and referred to the multi-disciplinary intervention, were eligible to participate in the present study if they had been diagnosed with GOA according to the definition proposed by Hoogeboom *et al.* [14]. Individuals were excluded from participation in the intervention if: 1) they were awaiting joint replacement surgery, 2) they had already participated unsuccessfully in a self-management programme for their GOA complaints, 3) their therapists suspected that they were suffering high levels of distress, 4) they did not master the Dutch language, or 5) they were illiterate. Recruitment and treatment of patients took place at the rheumatology outpatient department at the Maartenskliniek Woerden (the Netherlands).

The study protocol was reviewed and approved by the Institutional Review Board of the University Medical Centre Nijmegen (protocol number 2009/173), and did not fall within the remit of the Medical Research Involving Human Subjects Act.

Design

A randomised concurrent multiple-baseline single-case design was applied [13]. Participants completed repeated measurements during a baseline phase (phase A), a therapy-phase (phase B, 12 weeks) and a post-therapy phase (phase A'). Phase A acted as a control and was therefore compared with phases B and A'. By applying multiple baselines of varying length, observed effects of the treatment can be distinguished from effects due to chance [12,15,16], thus increasing internal validity. The total duration of phase A and A' was set at 7 weeks for each participant, and consequently participants with a longer phase A had a shorter phase A'. Participants were randomly assigned to a baseline and post-therapy period of either 2 and 5

weeks, 2.5 and 4.5 weeks, 3 and 4 weeks, …, or 5 and 2 weeks, respectively, using the Wampold-Worsham method [17] to increase statistical power. During the total study period of 19 weeks, participants completed diary measures twice a week, resulting in a total of 38 measurement points (14 during phase A and A' and 24 during phase B). Each diary measure comprised 14 VAS scales.

Measurements

Feasibility of the intervention

To evaluate the feasibility of the intervention, we assessed: 1) number of, and reasons for, drop-out during the intervention; 2) adherence to the intervention; 3) occurrence of adverse events related to the intervention; 4) participants' satisfaction with the intervention (straightforward question ranging from 0 (totally dissatisfied) to 10 (totally satisfied)); and 5) participants' satisfaction with the assessment procedure (straightforward yes/no questions).

Diary Measures

Diary measures comprised 14 VAS scales (scoring range from 0 to 10). Pain and fatigue were measured by single straightforward questions. Furthermore, 12 items derived from validated questionnaires were scored on a VAS scale. Kinesiophobia was measured with four VAS scales [18]. Self-efficacy was assessed using two questions from the Arthritis Self-Efficacy Scale [19]. Acceptance of the disease was measured with two questions from the subscale Acceptance of the Illness Cognition questionnaire [20], and illness perceptions were evaluated by two questions from the Illness Perception questionnaire [21]. To assess the specific complaints of each participant, we used the Patient-Specific Complaints questionnaire (PSK)) [22]. The most important complaint was assessed through the diary measure. For all scales, a

higher score represented unfavourable outcomes. Pain and self-efficacy were our primary outcome measures.

Pre- and post-intervention measures

At baseline, we collected data on age, sex, level of education, and duration of symptoms. Prior to the start of the programme, we also assessed participant's expectations about its effectiveness on a scale from 0 to 10 (0 representing 'No expectations whatsoever"). Pre- and post-intervention measures consisted of a set of validated questionnaires. We measured fatigue with the "Subjective Fatigue" subscale of the Checklist Individual Strength (CIS) [23], on which higher scores represent greater fatigue. Self-efficacy was evaluated with the General Self-Efficacy Scale [24], where higher scores represent higher levels of self-efficacy. Acceptance and helplessness were measured using the Illness Cognitions Questionnaire (ICQ) [25], where higher scores reflect higher levels of agreement with that generic illness cognition. As no specific questionnaires are available to assess the self-reported functional status of individuals with GOA, we used generic questionnaires for both the lower and upper extremities, namely the Lower Extremity Functional Scale (LEFS) [26] and the Disability of Arm, Shoulder and Hand (DASH), respectively [27]. Higher scores on the LEFS and DASH represent lower and greater disability, respectively.

Intervention

The group-based intervention (8 persons per group) lasted 12 weeks, comprised 10 sessions of approximately 1.5 hours per session, and was provided by an occupational therapist and physical therapist. The intervention aimed to increase the participants' knowledge of the disease, to optimise the participants' current lifestyle, and to enhance the participants' self-efficacy in controlling the disease. All participants received information on the disease and

how to manage the disease (i.e., recommendations on activity pacing, medication use, physical activity and weight reduction). To enhance the participants' self-efficacy, the 5-As model of behaviour change counselling was used [28]. During each session the individual goals were monitored and discussed. Moreover, participants were enrolled in a therapeutic activity programme to improve the quality of movement. Finally, participants were familiarised with different kinds of sports, tailored to the participants' complaints to prevent overexertion (i.e. tai chi, brisk walking, and therapeutic fitness). An overview of the intervention is depicted in Box 1. Participants were advised to implement these recommendations in their home situation.

Data analysis

All data were entered into the data-entry program Epidata [29]. Ten per cent of the data was entered twice to establish the quality of data entry. Missing data were described.

Diary data were analysed using the 2-Standardised Deviation (SD) band method [16] (visual inspection) and randomisation tests [30]. The 2-SD band was calculated from the baseline data and graphed from the baseline phase through the intervention phase. If two or more successive data points in the intervention or post-intervention phase fell outside the bandwidth of 2 SDs, the result was considered significant [16]. As serial dependence - the extent to which scores at one point in a series are predictive of scores at another point in the same data set - can bias the visual inspection [16], we checked our data in each phase for serial dependence using the lag-1 method [12]. If data were found to be significantly correlated, we transformed the data using a moving-average transformation, in which the preceding and succeeding measurements were taken into account [12,15]. In addition, randomisation tests for multiple-baseline single-case designs were carried out. We expected phase B and A' to be superior to phase A in terms of our health outcome assessment. Therefore our we tested the

null hypothesis - that there would be no differential effect for any of the measurement times - using a randomisation test of the differences in the means between the pre-intervention phase and the intervention or post-intervention phase [16]. A p-value < 0.05 was considered statistically significant. For the pre- and post-measurements, we considered change scores of 20% on validated questionnaires as clinically relevant. We used Stata/IC 10.1 for Windows for the descriptive and visual analysis of the data and R version 2.14.1 for the randomisation tests [30].

Results

Five participants gave written informed consent to participate in the study. One patient dropped out of the study within two weeks after the start of the study, reporting that filling out the questionnaires was too demanding for her on an emotional level. However, she did continue with the multi-disciplinary intervention. The four remaining participants completed all 38 diary measures, resulting in 2,128 completed items. Six items (0.3%) were missing. Data entry errors were negligible (<0.1%). Table 1 presents the characteristics of the participants.

Feasibility of the intervention

Prior to the intervention, participants' expectations regarding the effectiveness of the intervention ranged from 5 to 7 (median = 7). Participant 3 missed three of the 10 sessions; participants 2 and 4 both missed one session. Participant 1 reported an increase in pain levels, which she ascribed to the intervention. Satisfaction with the intervention was assigned a score of 8 points out of 10 by participants 1, 2 and 4, and 7 points out of 10 by participant 3. Perceived therapy effects were assigned a score of 7, 3, 5, and 7 out of 10 by participants 1, 2, 3 and 4, respectively. All participants believed the questionnaires used in this study properly evaluated their most important issues. The remarks most frequently made by participants regarding the intervention were: 1) there were too many sessions and these were too short/brief; 2) too much verbal information; 3) too much time between two sessions; 4) too little information on acceptance of the disease; and 5) too little individualisation in the exercise sessions, and in setting and monitoring therapy goals.

Diary measures

Our primary outcome measures were pain and self-efficacy. In the pain data, participant 3's intervention phase showed serial dependence, and that of participants 1 and 4 showed large fluctuations. Thus, we transformed these data prior to completion of visual data analysis. The 2-SD band method showed that participants 1, 2 and 4 each experienced significant deterioration in their pain scores between baseline, intervention and post-intervention phases. Participant 3 demonstrated significant improvement during the intervention phase (Figure 1), though this did not persist during the post-intervention phase. For all four participants, randomisation tests demonstrated no significant changes in pain between the pre-intervention phase and the intervention/post-intervention phase (p=0.93). Serial dependence was found in participant 4's self-efficacy data, and these data were transformed prior to the analyses. The 2-SD band method demonstrated that participant 4 experienced significantly higher levels of self-efficacy in both the intervention and post-intervention phase compared to the baseline phase. No differences were found for participants 1, 2 and 3. Randomisation testing demonstrated no statistically significant difference between the phase prior to the intervention and the phases during and after the intervention (p=0.85). Randomisation tests for our secondary outcome measures are shown in Table 2.

Pre- and post-measurements

Table 3 depicts the clinically relevant changes from baseline for each of the four participants. None of the participants reported improvement in self-efficacy. Participant 1 experienced clinically relevant deterioration in self-efficacy, upper body function and kinesiophobia. Participant 4 reported improvements in fatigue levels, upper body function, kinesiophobia and acceptance. Both participants 2 and 3 remained stable.

Discussion

Our data suggest that the tailored, 12-week non-pharmacological intervention for patients with GOA was feasible in terms of adverse events, number of drop-outs and participation rate. On the other hand, the participants raised several critical points concerning the structure, content, and perceived benefits of the intervention. The latter was confirmed by visual inspection of the data and randomisation testing, as the intervention did not demonstrate clear-cut effects on health-related factors. Therefore, we believe the content and structure of the current intervention does not warrant further evaluation in a randomised clinical trial.

In view of the participants' remarks, we believe that the intervention should be more individually tailored. One of the remarks was that the therapeutic movement programme was not sufficiently individualised to address the participants' health problems. In a future non-pharmacological intervention, it might be of value to incorporate the results of the Patient-Specific Complaints instrument [22] in the therapeutic activity programme. Moreover, it was suggested that setting and achieving goals should be monitored more closely. To do so, participants should draw up action plans by completing goal-setting forms to formulate short-term goals, whilst being aware of potential limiting factors. In this way, personal goals could be monitored, discussed and adjusted, which in turn might increase the involvement and self-efficacy of the participants [16]. Finally, participants had relatively low treatment expectations regarding the intervention (highest score was 7 out of 10), implying that participants might have lacked an active role prior to the start of intervention. Motivation is considered one of the most important factors for the success of a self-management programme [31,32]. Therefore, to increase the effectiveness of a non-pharmacological intervention in patients with GOA, attention should be paid to participants' motivation prior to inclusion.

Furthermore, therapists could be trained in motivating and goal-setting techniques, for example motivational interviewing.

Several limitations should be taken into account when interpreting our data. First, we used a concurrent multiple-baseline single-case design to evaluate the intervention's preliminary effectiveness. This design is particularly successful in demonstrating immediate effects [33]. Since our intervention aimed to improve self-management in individuals with osteoarthritis, which is often considered challenging and time-consuming [9], our choice of study design might not be optimal, given the short evaluation period and the considerable length of the treatment programme. A second limitation was that all participants were in the same therapy group, possibly resulting in a negative group effect compromising any therapy effects. On the other hand, the traditional approach to multiple-baseline studies is for all participants to undergo treatment simultaneously [13]. This strategy is recommended as it improves internal validity, particularly in terms of history effects [34]. A third limitation, inherent to the design of the study, is that the study has lower external validity than randomised clinical trials, for which participants are usually selected to form a generalizable sample [35]. A final limitation of this study was its inability to test the feasibility of study logistics for a randomised clinical trial (for example, recruitment rate, drop-out rate, and issues concerning randomisation) [36].

As far as we know, we are the first to study a multidisciplinary intervention to improve self-management in people with GOA. Due to differences in study populations, our results cannot be compared with those of another study into the effect of a non-pharmacological intervention in patients with GOA after major joint replacement surgery [37]. It is remarkable that so little research is available given the relatively high prevalence of joint pain comorbidity in individuals with established osteoarthritis and its association with compromised health status [1,2].

Some consider single-case experimental designs as viable alternatives to large-scale randomised clinical trials [38,39], whereas others state the opposite [35,40]. Whilst using this design, we faced several (practical) constraints that potential users should be aware of. As yet, there is a plethora of analytical techniques for single-case data [30], with little or no consensus on the optimal way to analyse the data. In our study, we demonstrated a significant effect of our intervention on kinesiophobia using a randomisation test, whereas visual inspection showed only clear effects in one participant. Another practical consideration is that the design requires a substantial contribution from the participants. In the present study, one of the participants dropped out as she experienced additional psychological burden due to recurring questionnaires. It remains to be elucidated whether frequent assessment of health status as in the current study negatively, or perhaps positively, influences health outcomes. In our opinion, the multiple-baseline single-case study is a useful and valid alternative to the randomised pilot study, as it gives insight into the feasibility and preliminary effectiveness, allowing one to tailor the content and context of the intervention prior to conducting a randomised clinical trial. However, it should only be considered an alternative to a full-sized randomised clinical trial in rare diseases or in situations where a randomised clinical trial is unfeasible or unethical, due to the low external validity of the findings.

An interesting finding was the marked variability in VAS scores within participants on specific outcomes. For example, three participants reported fluctuations in pain scores of more than 4 points within a period of half a week (i.e., between two measurement points). Fluctuations in pain between two measurement points ranged from 0 to 7 points, frequently exceeding the thresholds for clinically relevant differences [41]. Such fluctuations indicate that pain in OA is far less stable than often believed and should perhaps be assessed far more frequently. As such variations are also likely to occur in randomised clinical trials, researchers

should consider assessing post-intervention health outcomes at repeated time points. These outcomes could then be averaged to obtain a more stable post-intervention point estimate.

In conclusion, health providers and researchers should be aware of the lack of studies on the effectiveness of non-pharmacological interventions for patients with GOA. In our study, although we systematically conceptualised our intervention according to the latest evidence [7-9] and in collaboration with several health care providers, both feasibility and effectiveness of the care programme are doubtful. Therefore, the current therapy programme does not warrant evaluation in a large randomised clinical trial, although data obtained in this multiple baseline study have highlighted several ways in which the therapy program could be optimized/improved.

Reference List

(1) Hoogeboom TJ, den Broeder AA, Swierstra BA, et al. Joint-pain comorbidity, health status, and medication use in hip and knee osteoarthritis: A cross-sectional study. Arthritis Care Res (Hoboken) 2012 Jan;**64**:54-8.

(2) Suri P, Morgenroth DC, Kwoh CK, et al. Low back pain and other musculoskeletal pain comorbidities in individuals with symptomatic osteoarthritis of the knee: data from the osteoarthritis initiative. Arthritis Care Res (Hoboken) 2010 Dec;**62**:1715-23.

(3) Forestier R, Francon A, Briole V, et al. Prevalence of generalized osteoarthritis in a population with knee osteoarthritis. Joint Bone Spine 2011 May;**78**:275-8.

(4) Gunther KP, Sturmer T, Sauerland S, et al. Prevalence of generalised osteoarthritis in patients with advanced hip and knee osteoarthritis: the Ulm Osteoarthritis Study. Ann Rheum Dis 1998 Dec;**57**:717-23.

(5) Conaghan P, Birrell F, Burke M, et al. Osteoarthritis: national clinical guideline for care and management in adults. London: Royal College of Physicians 2008.

(6) Stukstette MJPM, Hoogeboom TJ, de Ruiter R, et al. A multidisciplinary and multidimensional intervention for patients with hand osteoarthritis. Clinical Rehabilitation 2012.

(7) Zhang W, Moskowitz RW, Nuki G, et al. OARSI recommendations for the management of hip and knee osteoarthritis, part I: critical appraisal of existing treatment guidelines and systematic review of current research evidence. Osteoarthritis Cartilage 2007 Sep;**15**:981-1000.

(8) Zhang W, Moskowitz RW, Nuki G, et al. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. Osteoarthritis Cartilage 2008 Feb;**16**:137-62.

(9) Zhang W, Nuki G, Moskowitz RW, et al. OARSI recommendations for the management of hip and knee osteoarthritis: part III: Changes in evidence following systematic cumulative update of research published through January 2009. Osteoarthritis Cartilage 2010 Apr;**18**:476-99.

(10) Craig P, Dieppe P, Macintyre S, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008;**337**:a1655.

(11) Eldridge SM, Ashby D, Feder GS, et al. Lessons for cluster randomized trials in the twenty-first century: a systematic review of trials in primary care. Clin Trials 2004 Feb;**1**:80-90.

(12) Backman CL, Harris SR, Chisholm JA, et al. Single-subject research in rehabilitation: a review of studies using AB, withdrawal, multiple baseline, and alternating treatments designs. Arch Phys Med Rehabil 1997 Oct;**78**:1145-53.

(13) Christ T. Experimental control and threats to internal validity of concurrent and nonconcurrent multiple baseline designs. Psychology in the Schools 2007;**44**:451-9.

(14) Hoogeboom TJ, Stukstette MJ, de Bie RA, et al. Non-pharmacological care for patients with generalized osteoarthritis: design of a randomized clinical trial. BMC Musculoskelet Disord 2010;**11**:142.

(15) Nourbakhsh MR, Ottenbacher KJ. The statistical analysis of single-subject data: a comparative examination. Phys Ther 1994 Aug;**74**:768-76.

(16) Holtgrefe K, McCloy C, Rome L. Changes associated with a quota-based approach on a walking program for individuals with fibromyalgia. J Orthop Sports Phys Ther 2007 Dec;**37**:717-24.

(17) Ferron J, Sentovich C. Statistical Power of Randomization Tests Used With Multiple-Baseline Designs. The Journal of Experimental Education 2002;**70**:165-78.

(18) de Jong JR, Vlaeyen JW, de Gelder JM, et al. Pain-related fear, perceived harmfulness of activities, and functional limitations in complex regional pain syndrome type I. J Pain 2011 Dec;**12**:1209-18.

(19) Lorig K, Holman H. Arthritis Self-Efficacy Scales measure self-efficacy. Arthritis Care Res 1998 Jun;**11**:155-7.

(20) Evers AW, Kraaimaat FW, van LW, et al. Beyond unfavorable thinking: the illness cognition questionnaire for chronic diseases. J Consult Clin Psychol 2001 Dec;**69**:1026-36.

(21) Bijsterbosch J, Scharloo M, Visser AW, et al. Illness perceptions in patients with osteoarthritis: change over time and association with disability. Arthritis Rheum 2009 Aug 15;**61**:1054-61.

(22) Beurskens AJ, de Vet HC, Koke AJ, et al. A patient-specific approach for measuring functional status in low back pain. J Manipulative Physiol Ther 1999 Mar;**22**:144-8.

(23) Beurskens AJ, Bultmann U, Kant I, et al. Fatigue among working people: validity of a questionnaire measure. Occup Environ Med 2000 May;**57**:353-7.

(24) Luszczynska A, Scholz U, Schwarzer R. The general self-efficacy scale: multicultural validation studies. J Psychol 2005 Sep;**139**:439-57.

(25) Maas M, Taal E, van der Linden S, et al. A review of instruments to assess illness representations in patients with rheumatic diseases. Ann Rheum Dis 2009 Mar;**68**:305-9.

(26) Hoogeboom TJ. Translation and validation of the Dutch Lower Extremity Functional Scale. Submitted 2012.

(27) Beaton DE, Katz JN, Fossel AH, et al. Measuring the whole or the parts? Validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. J Hand Ther 2001 Apr;**14**:128-46.

(28) Glasgow RE, Emont S, Miller DC. Assessing delivery of the five 'As' for patient-centered counseling. Health Promot Int 2006 Sep;**21**:245-55.

(29) EpiData Data Entry, Data Management and basic Statistical Analysis System. Odense Denmark, EpidData Association, 2000-2008. [computer program]. 2012.

(30) Bulte I, Onghena P. Randomization tests for multiple-baseline designs: an extension of the SCRT-R package. Behav Res Methods 2009 May;**41**:477-85.

(31) DeWalt DA, Davis TC, Wallace AS, et al. Goal setting in diabetes self-management: taking the baby steps to success. Patient Educ Couns 2009 Nov;**77**:218-23.

(32) Teixeira PJ, Going SB, Houtkooper LB, et al. Exercise motivation, eating, and body image variables as predictors of weight control. Med Sci Sports Exerc 2006 Jan;**38**:179-88.

(33) Neuman SB. Single-Subject Experimental Design. In: Duke NK, Mallette MH, editors. Literacy Research Methodologies. 2nd edition ed. New York: The Guilford Press; 2011. p. 383.

(34) Carr JE. Recommendations for reporting multiple-baseline designs across participants. Behav Intervent 2005;**20**:219-24.

(35) Newcombe RG. Should the single subject design be regarded as a valid alternative to the randomised controlled trial? Postgrad Med J 2005 Sep;**81**:546-7.

(36) Thabane L, Ma J, Chu R, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol 2010;**10**:1.

(37) Bulthuis Y, Drossaers-Bakker KW, Taal E, et al. Arthritis patients show long-term benefits from 3 weeks intensive exercise training directly following hospital discharge. Rheumatology (Oxford) 2007 Nov;**46**:1712-7.

(38) Rizvi SL, Nock MK. Single-case experimental designs for the evaluation of treatments for self-injurious and suicidal behaviors. Suicide Life Threat Behav 2008 Oct;**38**:498-510.

(39) Janosky JE. Use of the single subject design for practice based primary care research. Postgrad Med J 2005;**81**:549-51.

(40) Guyatt GH, Keller JL, Jaeschke R, et al. The n-of-1 randomized controlled trial: clinical usefulness. Our three-year experience. Ann Intern Med 1990 Feb 15;**112**:293-9.

(41) Lee JS, Hobden E, Stiell IG, et al. Clinically important change in the visual analog scale after adequate pain control. Acad Emerg Med 2003 Oct;**10**:1128-30.

Box 1. Pat-plot of the multi-disciplinary intervention.

Timeline		Intervention
Pre-measurement		
Week 1	Part 1	a
	Part 2	b c
Week 2	Part 1	d e
	Part 2	b c
Week 3	Part 1	d 1
	Part 2	c 1 2
Week 4		e 1 2
Week 5		b d 1
Week 6		c 1 2
Week 7		e 1
Week 9		b c 1
Week 12		d 1 f
Post-measurement		

a	Introduction meeting.
b	Information on Pain and Medication use.
c	Information on Activity Pacing.
d	Information on the importance of Physical Activity.
e	Information on Weight Reduction.
1	Activity programme to improve quality of movement.
2	Sports activity.
f	Evaluation time point.

Table 1. Characteristics of the study participants.

Participant	Sex	Age (y)	Education	No. painful joint groups (0 - 11)	Baseline assignment (measurements)
1	F	76	Low	8	4
2	F	68	Medium	3	5
3	M	59	Low	11	7
4	F	56	High	5	6
5[†]	F	51	High	-	6

Abbreviations: F, female; M, male; No., number of.

[†] Dropped out.

Table 2. Randomisation tests for the diary measurement outcomes.

	Diary measures Randomisation tests*
Pain	$p = 0.93$
Fatigue	$p = 0.79$
Self-efficacy	$p = 0.85$
Patient-specific complaints	$p = 0.64$
Kinesiophobia	$p =$ **0.02**
Illness cognition	$p = 0.69$
Illness perception	$p = 0.60$

*Predefined expectation was that phase B would be smaller than phase A.

Table 3. Clinically relevant differences between baseline and post-intervention measurements.

	Fatigue		Self-efficacy		Function				Kinesiophobia		Illness Cognitions			
					Upper		Lower				Help		Accept	
	T0	T1	T0	T1	T0	T1	T0	T1	T0	T1	T0	T1	T0	T1
p1	42	39	<u>35</u>	<u>27</u>	<u>35</u>	<u>50</u>	44	47	43	50	11	11	12	12
p2	9	9	35	37	**18**	**13**	69	68	28	31	8	9	23	24
p3	**56**	**33**	35	30	<u>31</u>	<u>43</u>	38	41	57	53	13	14	**15**	**19**
p4	**34**	**27**	29	31	**44**	**32**	46	48	**48**	**34**	9	9	**11**	**14**

Bold = 20% improvement, Underlined = 20% deterioration.

Accept = Subscale Acceptance; Help = Subscale Helplessness.

Figure 1. Diary measures for pain with 2-SD band graph for baseline, intervention and post-intervention phases. Scores on the pain VAS range from 0 to 10; higher scores indicate higher levels of pain.

Figure 2. Diary measures for Self-Efficacy with 2-SD band graph for baseline, intervention and post-intervention phases. Scores on the pain VAS range from 0 to 10, higher scores indicating lower levels of self-efficacy.

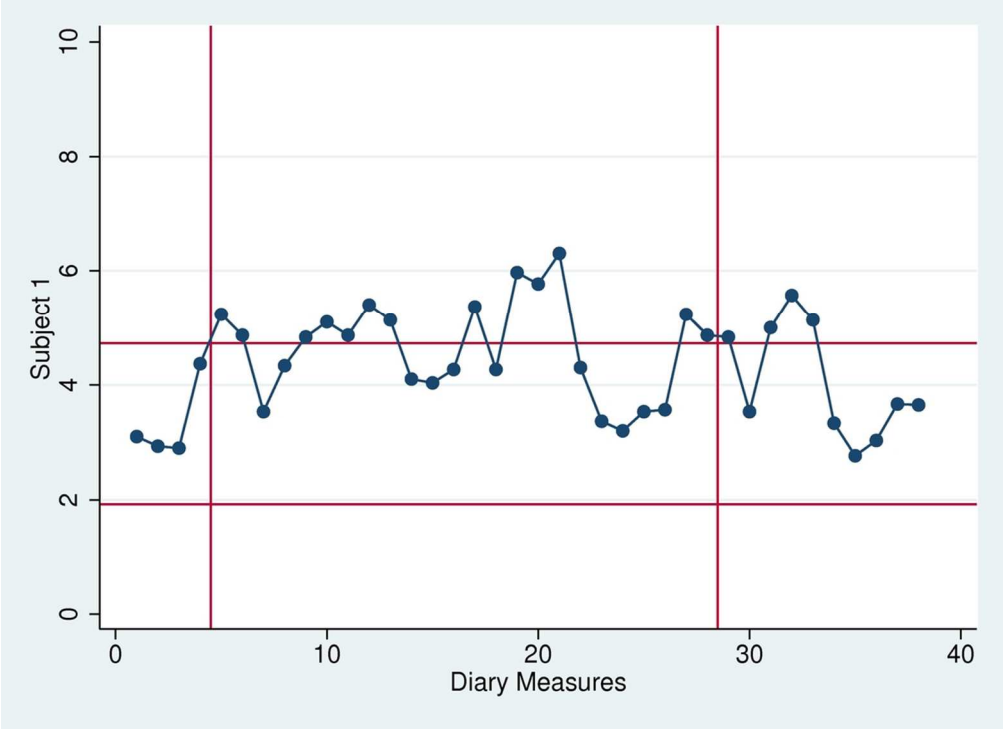


Figure 1. Diary measures for pain with 2-SD band graph for baseline, intervention and post-intervention phases. Scores on the pain VAS range from 0 to 10; higher scores indicate higher levels of pain.
101x73mm (300 x 300 DPI)

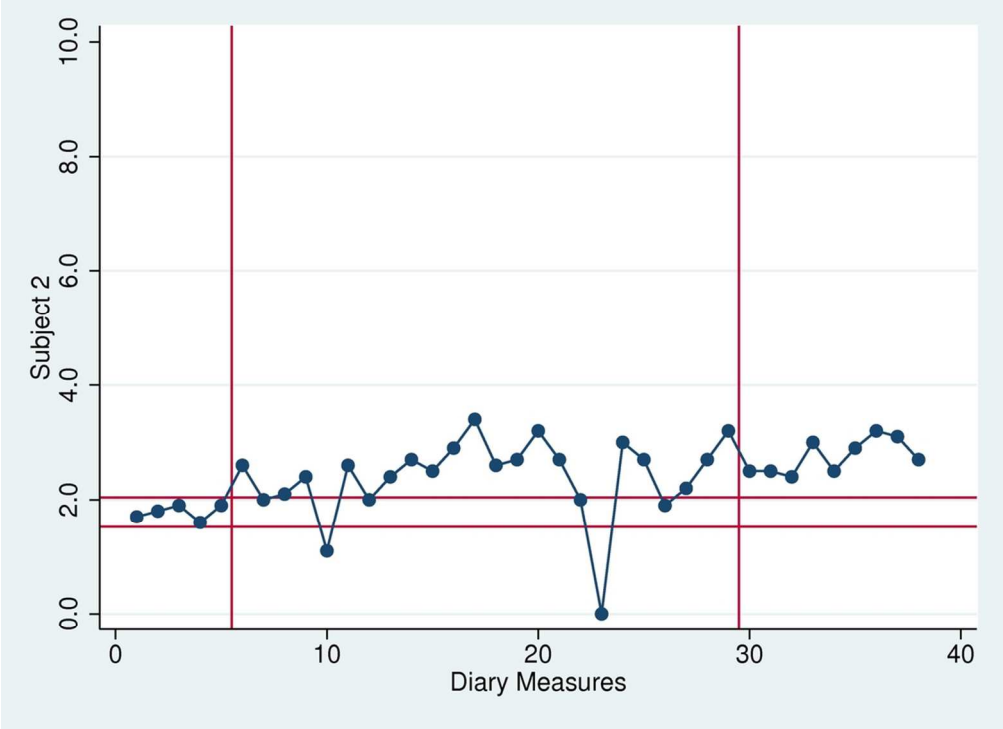


Figure 1. Diary measures for pain with 2-SD band graph for baseline, intervention and post-intervention phases. Scores on the pain VAS range from 0 to 10; higher scores indicate higher levels of pain.
101x73mm (300 x 300 DPI)

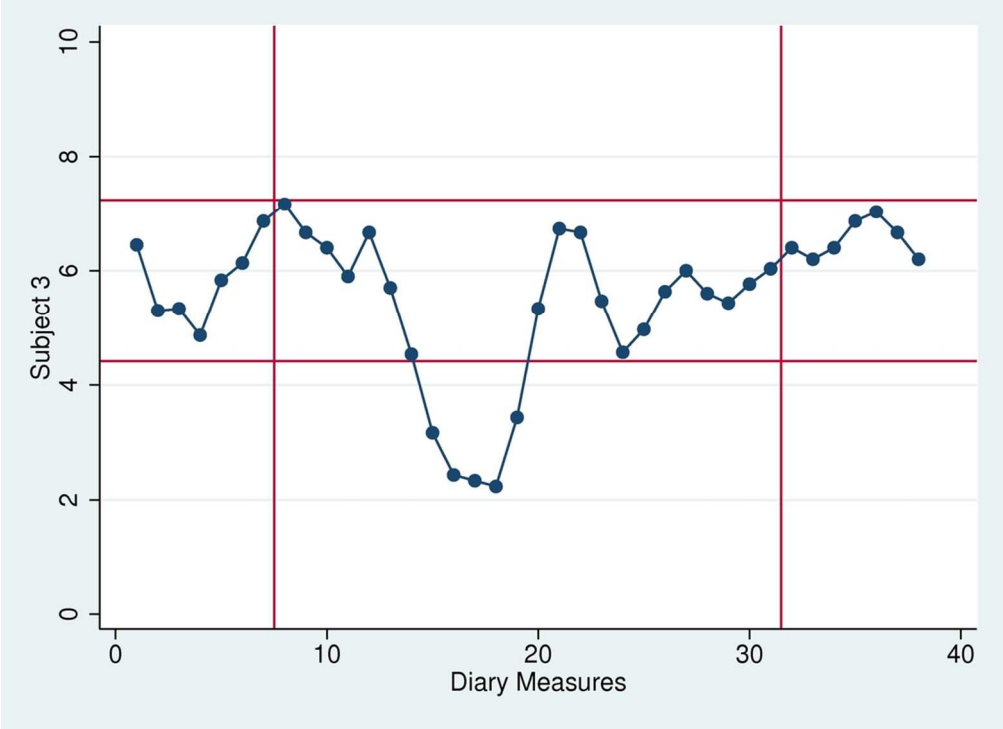


Figure 1. Diary measures for pain with 2-SD band graph for baseline, intervention and post-intervention phases. Scores on the pain VAS range from 0 to 10; higher scores indicate higher levels of pain.
101x73mm (300 x 300 DPI)

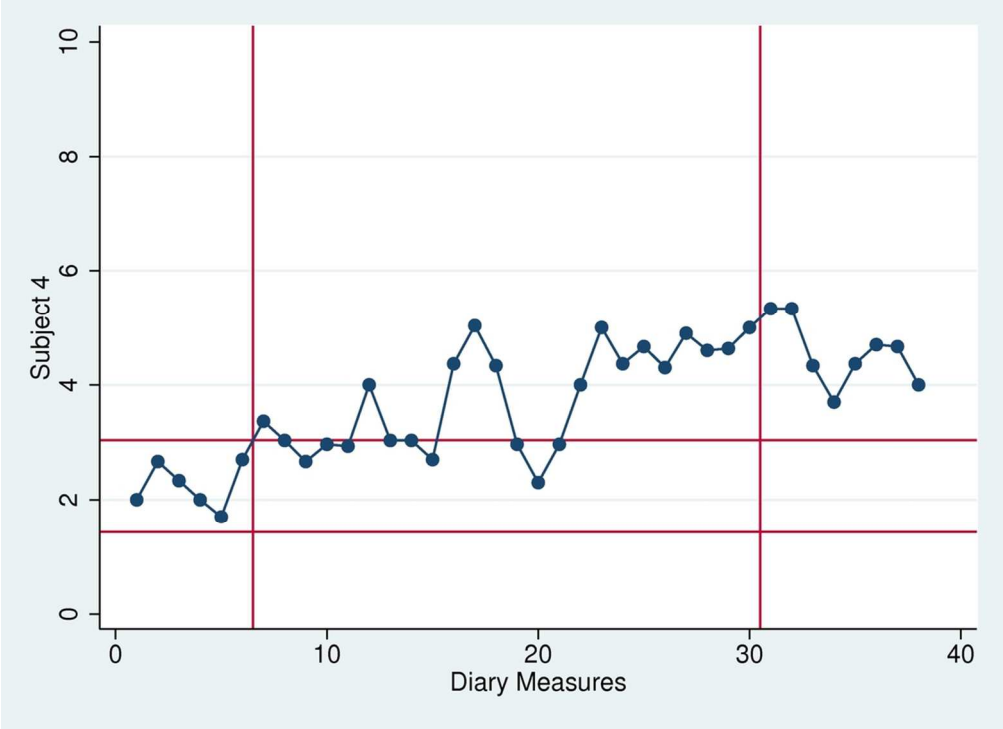


Figure 1. Diary measures for pain with 2-SD band graph for baseline, intervention and post-intervention phases. Scores on the pain VAS range from 0 to 10; higher scores indicate higher levels of pain.
101x73mm (300 x 300 DPI)

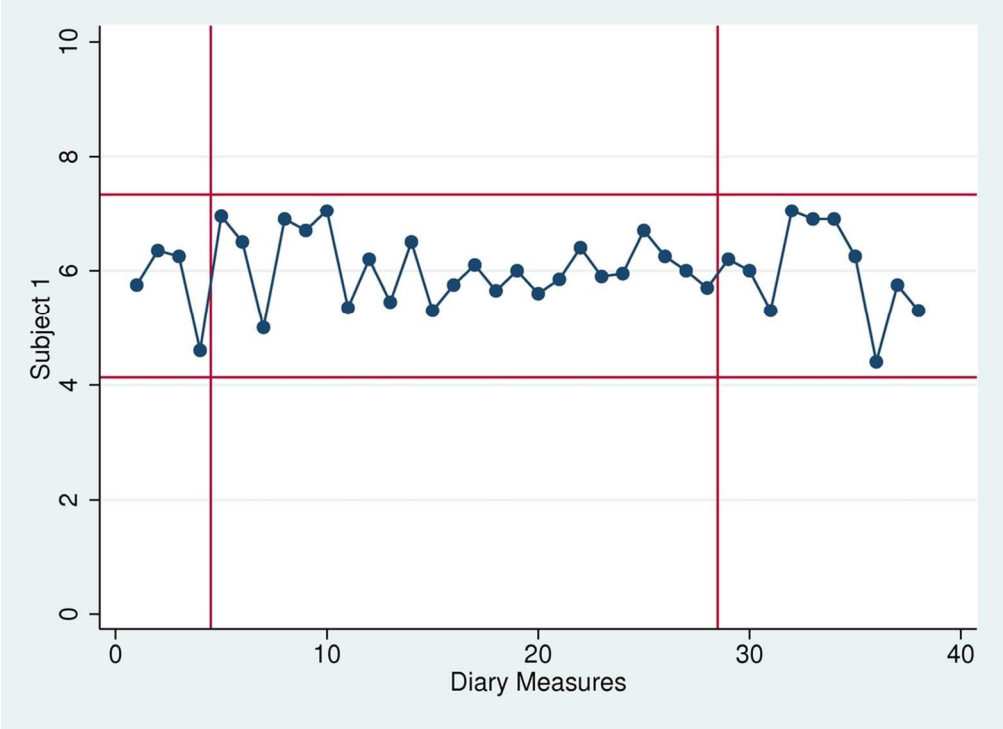


Figure 2. Diary measures for Self-Efficacy with 2-SD band graph for baseline, intervention and post-intervention phases. Scores on the pain VAS range from 0 to 10, higher scores indicating lower levels of self-efficacy.
101x73mm (300 x 300 DPI)

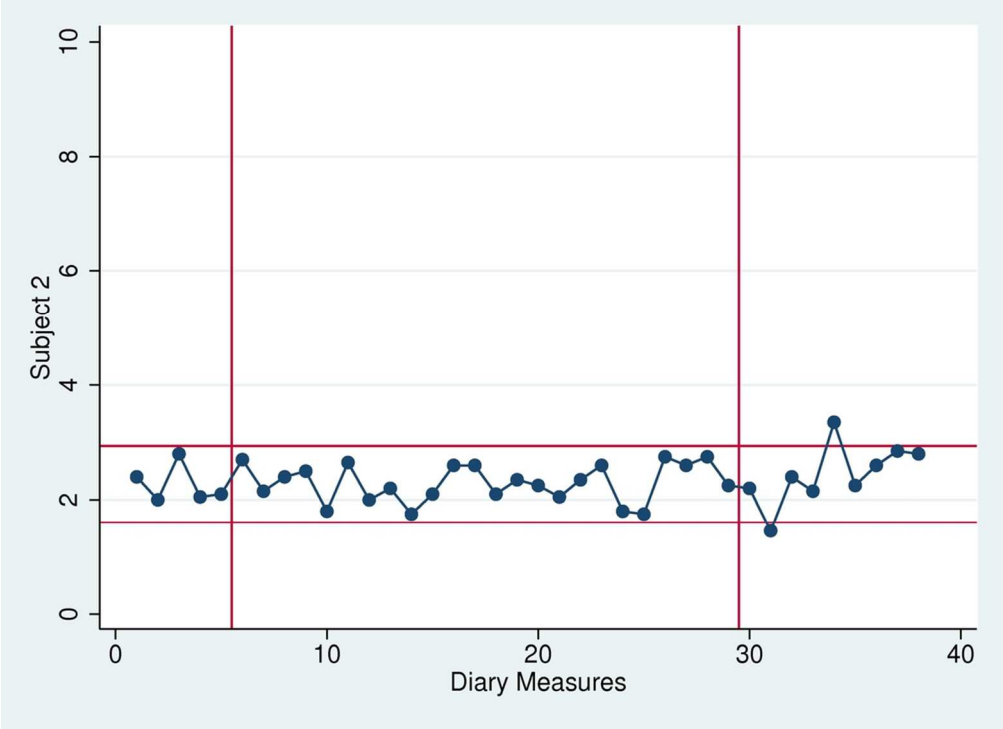
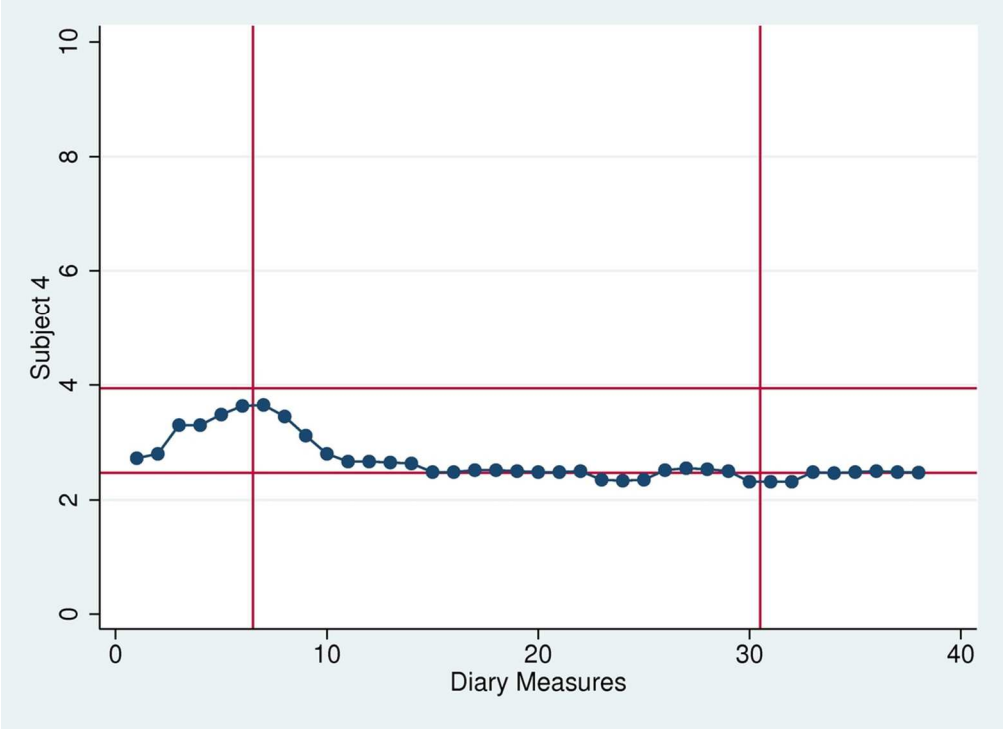
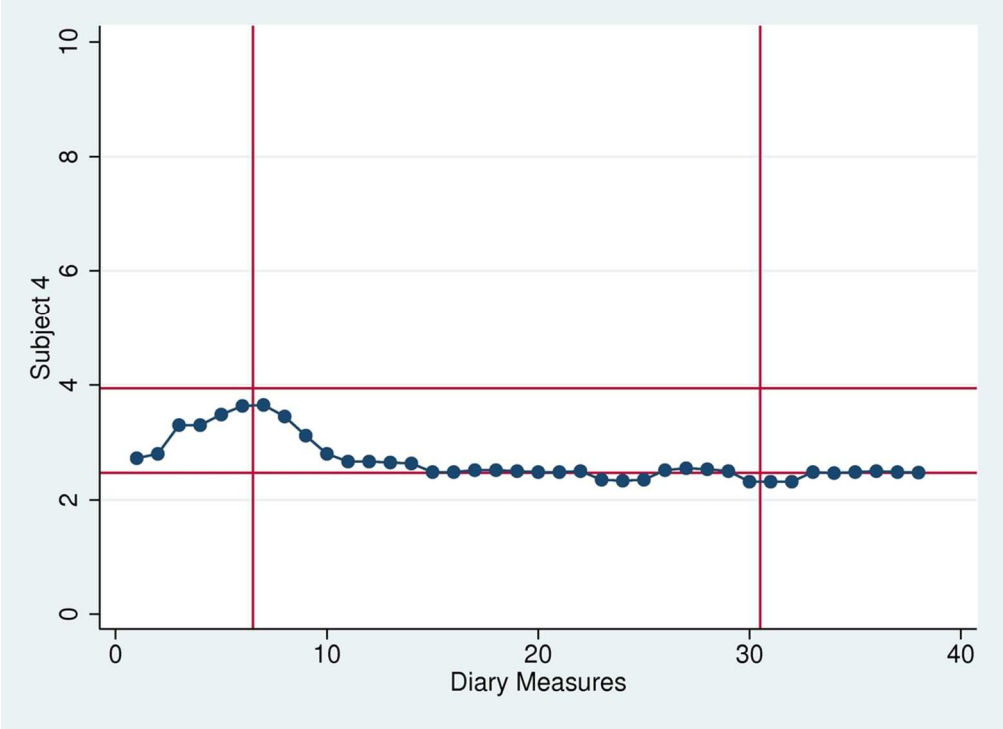


Figure 2. Diary measures for Self-Efficacy with 2-SD band graph for baseline, intervention and post-intervention phases. Scores on the pain VAS range from 0 to 10, higher scores indicating lower levels of self-efficacy.
101x73mm (300 x 300 DPI)



101x73mm (300 x 300 DPI)



101x73mm (300 x 300 DPI)

BMJ open

Feasibility and potential effectiveness of a non-pharmacological multidisciplinary care programme for persons with generalised osteoarthritis: a randomised multiple-baseline single-case study.

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2012-001161.R1
Article Type:	Research
Date Submitted by the Author:	14-Jun-2012
Complete List of Authors:	Hoogeboom, Thomas; Sint Maartenskliniek, Department of Rheumatology; Maastricht University Medical Centre, Department of Epidemiology Kwakkenbos, Linda; Sint Maartenskliniek, Department of Rheumatology Rietveld, Leonie; Maartenskliniek, Department of Rheumatology den Broeder, Alfons; Sint Maartenskliniek, Department of Rheumatology de Bie, Rob; Maastricht University Medical Centre, Department of Epidemiology van den Ende, Cornelia; Sint Maartenskliniek, Department of Rheumatology
Primary Subject Heading:	Rehabilitation medicine
Secondary Subject Heading:	Rehabilitation medicine, Evidence based practice, Pharmacology and therapeutics
Keywords:	Rehabilitation medicine < INTERNAL MEDICINE, Rheumatology < INTERNAL MEDICINE, Musculoskeletal disorders < ORTHOPAEDIC & TRAUMA SURGERY

Feasibility and potential effectiveness of a non-pharmacological multidisciplinary care programme for persons with generalised osteoarthritis: a randomised multiple-baseline single-case study.

Short title: Non-pharmacological care in generalized osteoarthritis

Hoogeboom TJ, Kwakkenbos L, Rietveld L, den Broeder AA, de Bie RA, van den Ende CHM.

§*Corresponding author:*

Tel: +31 24 327 2718,

Fax: +31 24 365 9154.

E-mail: t.hoogeboom@maartenskliniek.nl

Address for reprints requests:

Sint Maartenskliniek, ReumaResearch

PO Box 9011, 6500 GM, Nijmegen, the Netherlands.

Grant supporter: The study was financed by the Sint Maartenskliniek Nijmegen and Woerden, the Netherlands.

Competing interests: Authors have no competing interests to declare.

Manuscript word count: 3475 words.

Abstract

Objectives. To evaluate the feasibility and potential effectiveness of a 12-week non-pharmacological multidisciplinary intervention in patients with generalised osteoarthritis (GOA).

Design. A randomised, concurrent, multiple-baseline, single-case design. During the baseline period, the intervention period, and the post-intervention period, all participants completed several health outcomes twice a week on visual analogue scales.

Setting. Rheumatology, outpatient department of a specialized hospital in the Netherlands.

Participants. One man and four women (age 51 to 76) diagnosed with GOA.

Primary outcome measures. To assess feasibility we assessed the number of drop-outs and adverse events, adherence rates, and patient satisfaction.

Secondary outcome measures. To assess the potential effectiveness we assessed pain and self-efficacy using visual data inspection and randomisation tests.

Results. The intervention was feasible in terms of adverse events (none) and adherence rate, but not in terms of participant satisfaction with the intervention. Visual inspection of the data and randomisation testing demonstrated no effects on pain (p = 0.93) or self-efficacy (p = 0.85).

Conclusions. The results of the present study indicate that the proposed intervention for patients with GOA was insufficiently feasible and effective. The data obtained through this multiple-baseline study has highlighted several areas in which the therapy programme can be optimised.

Article summary

Article focus:

- To evaluate the feasibility the effectiveness of a 12-week non-pharmacological multidisciplinary intervention in patients with generalised osteoarthritis (GOA).
- To evaluate the potential effectiveness of a 12-week non-pharmacological multidisciplinary intervention in patients with GOA.

Key messages:

- To date no studies are available that evaluate non-pharmacological, multidisciplinary care in individuals with GOA.
- The intervention evaluated in the present study appeared both insufficiently feasible and effective for patients with GOA.
- Several areas in which the therapy programme could be optimised were highlighted.

Strengths and limitations of this study:

- A multiple-baseline single-case design is particularly successful in demonstrating immediate effects, whereas we studied changes in health behaviour.
- Inherent to the design of the study is lower external validity due to the small number of included participants.

Funding statement:

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors

Competing interests:

All authors declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work.

Author's contribution:

Substantial contribution to the conception and design of the study: TJH, LK, LR, AAB, RAB and CHME.

Substantial contribution to the acquisition of the data: TJH, LR

Substantial contribution to the analysis and interpretation of the data: TJH, LR, LK, CHME, AAB and CHME.

Provided intellectual content whilst drafting the article: TJH, LR, LK, CHME, AAB and CHME.

Approved the final version to be published: TJH, LR, LK, CHME, AAB and CHME.

Data sharing statement:

No additional data available.

Introduction

A growing body of evidence shows that individuals with established osteoarthritis with multiple joint involvement - often referred to as generalised osteoarthritis (GOA) - represent a relatively large subgroup of patients [1-4]. It has been suggested that these people might be in need of more intensive treatment options than patients with single joint complaints [1,5]. To the best of our knowledge, however, there are no studies that evaluate non-pharmacological, multidisciplinary care in individuals with GOA [5], warranting the development and evaluation of such a treatment programme. Therefore, we conceptualised a non-pharmacological, multidisciplinary treatment programme following a previously-described systematic procedure[6]. The intervention was based on recommendations for the management of hip and knee osteoarthritis [7-9], and was tailored to the needs of patients with multiple joint involvement [1]. Due to the complex nature of multiple joint-involvement in OA [1-4] and the fact that guidelines for hip and knee OA recommend multiple non-pharmacological treatment modalities, an intervention was developed by a multidisciplinary team [8].

Before evaluating such an intervention in a randomised clinical trial, a pilot study is recommended [10], since evaluations are often undermined by problems of acceptability, compliance, delivery of the intervention, recruitment and retention, and smaller-than-expected effect sizes [11]. A useful study design for pilot interventions is the multiple-baseline single-case design, as it allows researchers to test the feasibility of the intervention and to make an assessment of its potential effectiveness with a low number of participants [12]. In a multiple-baseline design, the intervention is introduced to subjects after randomly-assigned baseline periods of different lengths, and an effect is demonstrated if the measured outcome only changes after the intervention has been introduced [13].

The primary aim of our study was to evaluate the feasibility of a non-pharmacological, multidisciplinary intervention in patients with GOA. Our secondary aim was to assess the potential effectiveness of this intervention on pain and self-efficacy.

Methods

Participants

Men and women, 40 years or older and referred to the multidisciplinary intervention, were eligible to participate in the present study if they had been diagnosed with GOA; i.e. experiencing complaints in three or more joint groups, having at least two objective signs that indicate OA in at least two joints, and having limitations in daily functioning (Health Assessment Questionnaire-Disability Index score (HAQ-DI) [14] > 0.5) [15]. Individuals were excluded from participation in the intervention if: 1) they were awaiting joint replacement surgery, 2) they had already participated unsuccessfully in a self-management programme for their GOA complaints, 3) their therapists suspected that they were suffering high levels of distress, 4) they did not master the Dutch language, or 5) they were illiterate. Recruitment and treatment of patients took place at the rheumatology outpatient department at the Maartenskliniek Woerden (the Netherlands).

The study protocol was reviewed and approved by the Institutional Review Board of the University Medical Centre Nijmegen (protocol number 2009/173), and did not fall within the remit of the Medical Research Involving Human Subjects Act.

Design

A randomised concurrent multiple-baseline single-case design was applied [13]. Participants completed repeated measurements during a baseline phase (phase A), a therapy-phase (phase B, 12 weeks) and a post-therapy phase (phase A'). Phase A acted as a control and was therefore compared with phases B and A'. By applying multiple baselines of varying length, observed effects of the treatment can be distinguished from effects due to chance [12,16,17], thus increasing internal validity. The total duration of phase A and A' was set at 7 weeks for

each participant, and consequently participants with a longer phase A had a shorter phase A'. Participants were randomly assigned to a baseline and post-therapy period of either 2 and 5 weeks, 2.5 and 4.5 weeks, 3 and 4 weeks, …, or 5 and 2 weeks, respectively, using the Wampold-Worsham method [18] to increase statistical power. During the total study period of 19 weeks, participants completed diary measures twice a week, resulting in a total of 38 measurement points (14 during phase A and A' and 24 during phase B). Each diary measure comprised 14 VAS scales.

Measurements

Feasibility of the intervention

To evaluate the feasibility of the intervention, we assessed: 1) number of, and reasons for, drop-out during the intervention; 2) adherence to the intervention (number of no shows); 3) occurrence of adverse events related to the intervention; 4) participants' satisfaction with the intervention (straightforward question ranging from 0 (totally dissatisfied) to 10 (totally satisfied)); and 5) participants' satisfaction with the assessment procedure (straightforward yes/no questions).

Diary Measures

Diary measures comprised 14 VAS scales (scoring range from 0 to 10). Pain and fatigue were measured by single straightforward questions. Furthermore, 12 items derived from validated questionnaires were scored on a VAS scale. Kinesiophobia was measured with four VAS scales [19]. Self-efficacy was assessed using two questions from the Arthritis Self-Efficacy Scale [20]. Acceptance of the disease was measured with two questions from the subscale Acceptance of the Illness Cognition questionnaire [21], and illness perceptions were evaluated by two questions from the Illness Perception questionnaire [22]. To assess the specific

complaints of each participant, we used the Patient-Specific Complaints questionnaire (PSK)) [23]. The most important complaint was assessed through the diary measure. For all scales, a higher score represented unfavourable outcomes. Pain and self-efficacy were our main secondary outcome measures.

Pre- and post-intervention measures

At baseline, we collected data on age, sex, level of education (low (no or primary education), medium (secondary school and/or preparatory middle-level vocational education), high (university of applied sciences and/or university)) and duration of symptoms. Prior to the start of the programme, we also assessed participant's expectations about its effectiveness on a scale from 0 to 10 (0 representing 'No expectations whatsoever"). Pre- and post-intervention measures consisted of a set of validated questionnaires. We measured fatigue with the "Subjective Fatigue" subscale of the Checklist Individual Strength (CIS) [24], on which higher scores represent greater fatigue. Self-efficacy was evaluated with the General Self-Efficacy Scale [25], where higher scores represent higher levels of self-efficacy. Acceptance and helplessness were measured using the Illness Cognitions Questionnaire (ICQ) [26], where higher scores reflect higher levels of agreement with that generic illness cognition. As no specific questionnaires are available to assess the self-reported functional status of individuals with GOA, we used generic questionnaires for both the lower and upper extremities, namely the Lower Extremity Functional Scale (LEFS) [27] and the Disability of Arm, Shoulder and Hand (DASH), respectively [28]. Higher scores on the LEFS and DASH represent lower and greater disability, respectively.

Intervention

The group-based intervention (8 persons per group) lasted 12 weeks, comprised 10 sessions of approximately 1.5 hours per session, and was provided by an occupational therapist and physical therapist. To ensure group learning the treatment program was decided to be delivered in a group setting,. The intervention aimed to increase the participants' knowledge of the disease, to optimise the participants' current lifestyle, and to enhance the participants' self-efficacy in controlling the disease.

To do so, patients received information on activity pacing, medication use, physical activity and weight reduction. Consequently, based on the received information participants set personal goals regarding all these health areas. By setting these personal goals, participants transferred the health information into practical and personally relevant therapy goals. Goal setting and monitoring was done according to the 5-As model of behaviour change counselling [29]; a generally accepted method to enhance self-efficacy in health care settings. During each session, after the initial information session, the individual goals were monitored and discussed. To allow for positive feedback regarding the personal goals, all goals had to be achievable in brief amounts of time. Some examples of personal therapy goals were: 1. For the next three days, while at work, plan and perform 15 minutes of physical activity spread over three different time points (component Physical Activity); 2. For the next week, whilst cleaning the house, alternate (maximum of 10 minutes) between vacuum cleaning, other household chores, and rest moments (component Activity Pacing); 3. For the next week, use your pain medication (two tablets of Paracetamol (500 mg)) four times a day and monitor your pain during this period (component Medication Use); and 4. For the next week, eat at least three days two slices of whole wheat bread as breakfast (component Weight Reduction).

In addition, daily activities (such as walking, sitting, standing, stair climbing and getting in and out of bed) were included in the therapeutic activity programme. Participants received information and practised how to perform these daily activities without overexerting the joints

and muscles. Participants were instructed and encouraged to implement these techniques and methods of performing the activities in their daily practice.

Finally, participants were familiarised with different kinds of sports, tailored to the participants' complaints to prevent overexertion (i.e. tai chi, brisk walking, and therapeutic fitness). An overview of the intervention is depicted in Box 1. Participants were advised to implement these recommendations in their home situation.

Data analysis

All data were entered into the data-entry program Epidata [30]. Ten per cent of the data was entered twice to establish the quality of data entry. Missing data were described.

Diary data were analysed using the 2-Standardised Deviation (SD) band method [17] (visual inspection) and randomisation tests [31]. The 2-SD band was calculated from the baseline data and graphed from the baseline phase through the intervention phase. If two or more successive data points in the intervention or post-intervention phase fell outside the bandwidth of 2 SDs, the result was considered significant [17]. As serial dependence - the extent to which scores at one point in a series are predictive of scores at another point in the same data set - can bias the visual inspection [17], we checked our data in each phase for serial dependence using the lag-1 method [12]. If data were found to be significantly correlated, we transformed the data using a moving-average transformation, in which the preceding and succeeding measurements were taken into account [12,16]. In addition, randomisation tests for multiple-baseline single-case designs were carried out. We expected phase B and A' to be superior to phase A in terms of our health outcome assessment. Therefore our we tested the null hypothesis - that there would be no differential effect for any of the measurement times - using a randomisation test of the differences in the means between the pre-intervention phase and the intervention or post-intervention phase [17]. A p-value < 0.05 was considered

statistically significant. For the pre- and post-measurements, we considered change scores of 20% on validated questionnaires as clinically relevant [32]. We used Stata/IC 10.1 for Windows for the descriptive and visual analysis of the data and R version 2.14.1 for the randomisation tests [31].

Results

Nine people were screened to participate in the study; two patients were excluded as they did not report functional disabilities (HAQ-DI < 0.5) and two patients who were eligible were unable to attend the program. Eventually, five participants gave written informed consent to participate in the study. One patient dropped out of the study within two weeks after the start of the study, reporting that filling out the questionnaires was too demanding for her on an emotional level. However, she did continue with the multidisciplinary intervention. The four remaining participants completed all 38 diary measures, resulting in 2,128 completed items. Six items (0.3%) were missing. Data entry errors were negligible (<0.1%). Table 1 presents the characteristics of the participants.

Feasibility of the intervention

Prior to the intervention, participants' expectations regarding the effectiveness of the intervention ranged from 5 to 7 (median = 7). Participant 3 missed three of the 10 sessions; participants 2 and 4 both missed one session. Participant 1 reported an increase in pain levels, which she ascribed to the intervention. Satisfaction with the intervention was assigned a score of 8 points out of 10 by participants 1, 2 and 4, and 7 points out of 10 by participant 3. Perceived therapy effects were assigned a score of 7, 3, 5, and 7 out of 10 by participants 1, 2, 3 and 4, respectively. All participants believed the questionnaires used in this study properly evaluated their most important issues. The remarks most frequently made by participants regarding the intervention were: 1) there were too many sessions and these were too short/brief; 2) too much verbal information; 3) too much time between two sessions; 4) too little information on acceptance of the disease; and 5) too little individualisation in the exercise sessions, and in setting and monitoring therapy goals.

Diary measures

Our primary effectiveness outcome measures were pain and self-efficacy. In the pain data, participant 3's intervention phase showed serial dependence, and that of participants 1 and 4 showed large fluctuations. Thus, we transformed these data prior to completion of visual data analysis. The 2-SD band method showed that participants 1, 2 and 4 each experienced significant deterioration in their pain scores between baseline, intervention and post-intervention phases. Participant 3 demonstrated significant improvement during the intervention phase (Figure 1), though this did not persist during the post-intervention phase. For all four participants, randomisation tests demonstrated no significant changes in pain between the pre-intervention phase and the intervention/post-intervention phase (p=0.93). Serial dependence was found in participant 4's self-efficacy data, and these data were transformed prior to the analyses. The 2-SD band method demonstrated that participant 4 experienced significantly higher levels of self-efficacy in both the intervention and post-intervention phase compared to the baseline phase. No differences were found for participants 1, 2 and 3. Randomisation testing demonstrated no statistically significant difference between the phase prior to the intervention and the phases during and after the intervention (p=0.85). Outcomes of the randomisation tests for our secondary effectiveness outcome measures were: fatigue (p=0.79), patient specific complaints (p=0.64), kinesiophobia (p=0.02), illness cognitions (p=0.69) and illness perception (p=0.60).

Pre- and post-measurements

Table 2 depicts the clinically relevant changes from baseline for each of the four participants. None of the participants reported improvement in self-efficacy. Participant 1 experienced clinically relevant deterioration in self-efficacy, upper body function and kinesiophobia.

Participant 4 reported improvements in fatigue levels, upper body function, kinesiophobia and acceptance. Both participants 2 and 3 remained stable.

Discussion

Our data suggest that the tailored, 12-week non-pharmacological, multidisciplinary intervention for patients with GOA was feasible in terms of adverse events, number of drop-outs and participation rate. On the other hand, the participants raised several critical points concerning the structure, content, and perceived benefits of the intervention. The latter was confirmed by visual inspection of the data and randomisation testing, as the intervention did not demonstrate clear-cut effects on health-related factors. Therefore, we believe the content and structure of the current intervention does not warrant further evaluation in a randomised clinical trial.

In view of the participants' remarks, we believe that the intervention should be more individually tailored. One of the remarks was that the therapeutic movement programme was not sufficiently individualised to address the participants' health problems. In a future non-pharmacological, multidisciplinary intervention, it might be of value to incorporate the results of the Patient-Specific Complaints instrument [23] in the therapeutic activity programme. Moreover, it was suggested that setting and achieving goals should be monitored more closely. To do so, participants should draw up action plans by completing goal-setting forms to formulate short-term goals, whilst being aware of potential limiting factors. In this way, personal goals could be monitored, discussed and adjusted, which in turn might increase the involvement and self-efficacy of the participants [17]. Finally, participants had relatively low treatment expectations regarding the intervention (highest score was 7 out of 10), implying that participants might have lacked an active role prior to the start of intervention. Motivation is considered one of the most important factors for the success of a self-management programme [33,34]. Therefore, to increase the effectiveness of a non-pharmacological, multidisciplinary intervention in patients with GOA, attention should be paid to participants'

motivation prior to inclusion. Furthermore, therapists could be trained in motivating and goal-setting techniques, for example motivational interviewing.

Several limitations should be taken into account when interpreting our data. First, we used a concurrent multiple-baseline single-case design to evaluate the intervention's potential effectiveness. This design is particularly successful in demonstrating immediate effects [35]. Since our intervention aimed to improve self-management in individuals with osteoarthritis, which is often considered challenging and time-consuming [9], our choice of study design might not be optimal, given the short evaluation period and the considerable length of the treatment programme. A second limitation was that all participants were in the same therapy group, possibly resulting in a negative group effect compromising any therapy effects. On the other hand, the traditional approach to multiple-baseline studies is for all participants to undergo treatment simultaneously [13]. This strategy is recommended as it improves internal validity, particularly in terms of history effects [36]. A third limitation, inherent to the design of the study, is that the study has lower external validity than randomised clinical trials, for which participants are usually selected to form a generalizable sample [37]. A fourth limitation of this study was its inability to test the feasibility of study logistics for a randomised clinical trial (for example, recruitment rate, drop-out rate, and issues concerning randomisation) [38]. A final limitation was that we selected patients based on their medical diagnosis and functional status rather than on their scores on our main secondary outcomes (i.e. pain and/or self-efficacy). Future studies should include clinically relevant thresholds for their outcome measures in the in- and exclusion criteria.

As far as we know, we are the first to study a multidisciplinary intervention to improve self-management in people with GOA. Due to differences in study populations, our results cannot be compared with those of another study into the effect of a non-pharmacological, multidisciplinary intervention in patients with GOA after major joint replacement surgery

[39]. It is remarkable that so little research is available given the relatively high prevalence of individuals with established osteoarthritis with multiple joint involvement and its association with compromised health status [1,2].

Some consider single-case experimental designs as viable alternatives to large-scale randomised clinical trials [40,41], whereas others state the opposite [37,42]. Whilst using this design, we faced several (practical) constraints that potential users should be aware of. As yet, there is a plethora of analytical techniques for single-case data [31], with little or no consensus on the optimal way to analyse the data. In our study, we demonstrated a significant effect of our intervention on kinesiophobia using a randomisation test, whereas visual inspection showed only clear effects in one participant. Another practical consideration is that the design requires a substantial contribution from the participants. In the present study, one of the participants dropped out as she experienced additional psychological burden due to recurring questionnaires. It remains to be elucidated whether frequent assessment of health status as in the current study negatively, or perhaps positively, influences health outcomes. In our opinion, the multiple-baseline single-case study is a useful and valid alternative to the randomised pilot study, as it gives insight into the feasibility of the intervention and allows to evaluate the intervention's potential effectiveness, allowing one to tailor the content and context of the intervention prior to conducting a randomised clinical trial. However, single-case studies should only be considered an alternative to a full-sized randomised clinical trial in rare diseases or in situations where a randomised clinical trial is unfeasible or unethical, because of the designs' limitations, including low external validity of the findings and the inability to correct for confounders (such as medication use, age, disease duration etc.).

An interesting finding was the marked variability in VAS scores within participants on specific outcomes. For example, three participants reported fluctuations in pain scores of more than 4 points within a period of half a week (i.e., between two measurement points).

Fluctuations in pain between two measurement points ranged from 0 to 7 points, frequently exceeding the thresholds for clinically relevant differences [43]. Such fluctuations indicate that pain in OA is far less stable than often believed and should perhaps be assessed far more frequently. As such variations are also likely to occur in randomised clinical trials, researchers should consider assessing post-intervention health outcomes at repeated time points. These outcomes could then be averaged to obtain a more stable post-intervention point estimate.

In conclusion, health providers and researchers should be aware of the lack of studies on the effectiveness of non-pharmacological and/or multidisciplinary interventions for patients with GOA. In our study, although we systematically conceptualised our intervention according to the latest evidence [7-9] and in collaboration with several health care providers, both feasibility and effectiveness of the care programme are doubtful. Therefore, the therapy programme as described in this paper does not warrant evaluation in a large randomised clinical trial. Since the data obtained in this multiple baseline study have highlighted several ways in which the therapy program could be optimized/improved, these changes should be implemented prior to conducting an RCT to further examine the interventions' effectiveness.

Reference List

1 Hoogeboom TJ, den Broeder AA, Swierstra BA, et al. Joint-pain comorbidity, health status, and medication use in hip and knee osteoarthritis: A cross-sectional study. Arthritis Care Res (Hoboken) 2012;**64**:54-8.

2 Suri P, Morgenroth DC, Kwoh CK, et al. Low back pain and other musculoskeletal pain comorbidities in individuals with symptomatic osteoarthritis of the knee: data from the osteoarthritis initiative. Arthritis Care Res (Hoboken) 2010;**62**:1715-23.

3 Forestier R, Francon A, Briole V, et al. Prevalence of generalized osteoarthritis in a population with knee osteoarthritis. Joint Bone Spine 2011;**78**:275-8.

4 Gunther KP, Sturmer T, Sauerland S, et al. Prevalence of generalised osteoarthritis in patients with advanced hip and knee osteoarthritis: the Ulm Osteoarthritis Study. Ann Rheum Dis 1998;**57**:717-23.

5 Conaghan P, Birrell F, Burke M, et al. Osteoarthritis: national clinical guideline for care and management in adults. London: Royal College of Physicians 2008.

6 Stukstette M, Hoogeboom T, de RR, et al. A multidisciplinary and multidimensional intervention for patients with hand osteoarthritis. Clin Rehabil 2012;**26**:99-110.

7 Zhang W, Moskowitz RW, Nuki G, et al. OARSI recommendations for the management of hip and knee osteoarthritis, part I: critical appraisal of existing treatment guidelines and systematic review of current research evidence. Osteoarthritis Cartilage 2007;**15**:981-1000.

8 Zhang W, Moskowitz RW, Nuki G, et al. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. Osteoarthritis Cartilage 2008;**16**:137-62.

9 Zhang W, Nuki G, Moskowitz RW, et al. OARSI recommendations for the management of hip and knee osteoarthritis: part III: Changes in evidence following systematic cumulative update of research published through January 2009. Osteoarthritis Cartilage 2010;**18**:476-99.

10 Craig P, Dieppe P, Macintyre S, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008;**337**:a1655.

11 Eldridge SM, Ashby D, Feder GS, et al. Lessons for cluster randomized trials in the twenty-first century: a systematic review of trials in primary care. Clin Trials 2004;**1**:80-90.

12 Backman CL, Harris SR, Chisholm JA, et al. Single-subject research in rehabilitation: a review of studies using AB, withdrawal, multiple baseline, and alternating treatments designs. Arch Phys Med Rehabil 1997;**78**:1145-53.

13 Christ T. Experimental control and threats to internal validity of concurrent and nonconcurrent multiple baseline designs. Psychology in the Schools 2007;**44**:451-9.

14 Fries JF, Spitz P, Kraines RG, et al. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;**23**:137-45.

15 Hoogeboom TJ, Stukstette MJ, de Bie RA, et al. Non-pharmacological care for patients with generalized osteoarthritis: design of a randomized clinical trial. BMC Musculoskelet Disord 2010;**11**:142.

16 Nourbakhsh MR, Ottenbacher KJ. The statistical analysis of single-subject data: a comparative examination. Phys Ther 1994;**74**:768-76.

17 Holtgrefe K, McCloy C, Rome L. Changes associated with a quota-based approach on a walking program for individuals with fibromyalgia. J Orthop Sports Phys Ther 2007;**37**:717-24.

18 Ferron J, Sentovich C. Statistical Power of Randomization Tests Used With Multiple-Baseline Designs. The Journal of Experimental Education 2002;**70**:165-78.

19 de Jong JR, Vlaeyen JW, de Gelder JM, et al. Pain-related fear, perceived harmfulness of activities, and functional limitations in complex regional pain syndrome type I. J Pain 2011;**12**:1209-18.

20 Lorig K, Holman H. Arthritis Self-Efficacy Scales measure self-efficacy. Arthritis Care Res 1998;**11**:155-7.

21 Evers AW, Kraaimaat FW, van LW, et al. Beyond unfavorable thinking: the illness cognition questionnaire for chronic diseases. J Consult Clin Psychol 2001;**69**:1026-36.

22 Bijsterbosch J, Scharloo M, Visser AW, et al. Illness perceptions in patients with osteoarthritis: change over time and association with disability. Arthritis Rheum 2009;**61**:1054-61.

23 Beurskens AJ, de Vet HC, Koke AJ, et al. A patient-specific approach for measuring functional status in low back pain. J Manipulative Physiol Ther 1999;**22**:144-8.

24 Beurskens AJ, Bultmann U, Kant I, et al. Fatigue among working people: validity of a questionnaire measure. Occup Environ Med 2000;**57**:353-7.

25 Luszczynska A, Scholz U, Schwarzer R. The general self-efficacy scale: multicultural validation studies. J Psychol 2005;**139**:439-57.

26 Maas M, Taal E, van der Linden S, et al. A review of instruments to assess illness representations in patients with rheumatic diseases. Ann Rheum Dis 2009;**68**:305-9.

27 Hoogeboom TJ. Translation and validation of the Dutch Lower Extremity Functional Scale. Submitted 2012.

28 Beaton DE, Katz JN, Fossel AH, et al. Measuring the whole or the parts? Validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. J Hand Ther 2001;**14**:128-46.

29 Glasgow RE, Emont S, Miller DC. Assessing delivery of the five 'As' for patient-centered counseling. Health Promot Int 2006;**21**:245-55.

30 EpiData Data Entry, Data Management and basic Statistical Analysis System. Odense Denmark, EpidData Association, 2000-2008. [computer program]. 2012.

31 Bulte I, Onghena P. Randomization tests for multiple-baseline designs: an extension of the SCRT-R package. Behav Res Methods 2009;**41**:477-85.

32 Escobar A, Gonzalez M, Quintana JM, et al. Patient acceptable symptom state and OMERACT-OARSI set of responder criteria in joint replacement. Identification of cut-off values. Osteoarthritis Cartilage 2012;**20**:87-92.

33 DeWalt DA, Davis TC, Wallace AS, et al. Goal setting in diabetes self-management: taking the baby steps to success. Patient Educ Couns 2009;**77**:218-23.

34 Teixeira PJ, Going SB, Houtkooper LB, et al. Exercise motivation, eating, and body image variables as predictors of weight control. Med Sci Sports Exerc 2006;**38**:179-88.

35 Neuman SB. Single-Subject Experimental Design. In: Duke NK, Mallette MH, editors. Literacy Research Methodologies. 2nd edition ed. New York: The Guilford Press; 2011. p. 383.

36 Carr JE. Recommendations for reporting multiple-baseline designs across participants. Behav Intervent 2005;**20**:219-24.

37 Newcombe RG. Should the single subject design be regarded as a valid alternative to the randomised controlled trial? Postgrad Med J 2005;**81**:546-7.

38 Thabane L, Ma J, Chu R, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol 2010;**10**:1.

39 Bulthuis Y, Drossaers-Bakker KW, Taal E, et al. Arthritis patients show long-term benefits from 3 weeks intensive exercise training directly following hospital discharge. Rheumatology (Oxford) 2007;**46**:1712-7.

40 Rizvi SL, Nock MK. Single-case experimental designs for the evaluation of treatments for self-injurious and suicidal behaviors. Suicide Life Threat Behav 2008;**38**:498-510.

41 Janosky JE. Use of the single subject design for practice based primary care research. Postgrad Med J 2005;**81**:549-51.

42 Guyatt GH, Keller JL, Jaeschke R, et al. The n-of-1 randomized controlled trial: clinical usefulness. Our three-year experience. Ann Intern Med 1990;**112**:293-9.

43 Lee JS, Hobden E, Stiell IG, et al. Clinically important change in the visual analog scale after adequate pain control. Acad Emerg Med 2003;**10**:1128-30.

Box 1. Pat-plot of the multidisciplinary intervention.

Timeline		Intervention
Pre-measurement		
Week 1	Part 1	a
	Part 2	b c
Week 2	Part 1	d e
	Part 2	b c
Week 3	Part 1	d 1
	Part 2	c 1 2
Week 4		e 1 2
Week 5		b d 1
Week 6		c 1 2
Week 7		e 1
Week 9		b c 1
Week 12		d 1 f
Post-measurement		

a	Introduction meeting.
b	Information on Pain and Medication use.
c	Information on Activity Pacing.
d	Information on the importance of Physical Activity.
e	Information on Weight Reduction.
1	Activity programme to improve quality of movement.
2	Sports activity.
f	Evaluation time point.

1

Table 1. Characteristics of the study participants.

Participant	Sex	Age (y)	Education	No. painful joint groups (0 - 11)	Baseline assignment (measurements)
1	F	76	Low	8	4
2	F	68	Medium	3	5
3	M	59	Low	11	7
4	F	56	High	5	6
5†	F	51	High	-	6

Abbreviations: F, female; M, male; No., number of.

† Dropped out.

2

3

1

Table 2. Clinically relevant differences between baseline and post-intervention measurements.

	Fatigue		Self-efficacy		Function				Kinesiophobia		Illness Cognitions			
					Upper		Lower				Help		Accept	
	T0	T1	T0	T1	T0	T1	T0	T1	T0	T1	T0	T1	T0	T1
pt1	42	39	35	27	35	50	44	47	43	50	11	11	12	12
pt2	9	9	35	37	**18**	**13**	69	68	28	31	8	9	23	24
pt3	**56**	**33**	35	30	31	43	38	41	57	53	13	14	**15**	**19**
pt4	**34**	**27**	29	31	**44**	**32**	46	48	**48**	**34**	9	9	**11**	**14**

Bold = 20% improvement, Underlined = 20% deterioration. Abbreviations: Accept = Subscale Acceptance; Help = Subscale Helplessness; Lower = Lower extremity functioning; pt# = Participant #; T0 = Baseline measurement; T1 = Post-intervention measurement; Upper = Upper extremity functioning.

2

3

4

Figure 1. Diary measures for pain with 2-SD horizontal band graph for baseline (phase A), intervention (phase B) and post-intervention (phase A') phases. Scores on the pain VAS range from 0 to 10; higher scores indicate higher levels of pain.

Figure 2. Diary measures for Self-Efficacy with 2-SD horizontal band graph for baseline (phase A), intervention (phase B) and post-intervention (phase A') phases. Scores on the pain VAS range from 0 to 10, higher scores indicating lower levels of self-efficacy.

Feasibility and ~~preliminary e~~potential effectiveness ~~on pain and self-efficacy~~ of a non-pharmacological multidisciplinary care programme ~~for~~ for persons with generalised osteoarthritis: a ~~concurrent~~ randomised multiple-baseline single-case study.

Short title: Non-pharmacological care in generalized osteoarthritis

Hoogeboom TJ, Kwakkenbos L, Rietveld L, den Broeder AA, de Bie RA, van den Ende CHM.

§*Corresponding author:*

Tel: +31 24 327 2718,

Fax: +31 24 365 9154.

E-mail: t.hoogeboom@maartenskliniek.nl

Address for reprints requests:

Sint Maartenskliniek, ReumaResearch

PO Box 9011, 6500 GM, Nijmegen, the Netherlands.

Grant supporter: The study was financed by the Sint Maartenskliniek Nijmegen and Woerden, the Netherlands.

Competing interests: Authors have no competing interests to declare.

Manuscript word count: ~~2999~~ 3475 words.

Abstract

Objectives. To evaluate the feasibility and ~~to preliminarily evaluate the~~potential effectiveness of a 12-week ~~multidisciplinary~~ non-pharmacological multidisciplinary intervention in patients with generalised osteoarthritis (GOA).

Design. A randomised, concurrent, multiple-baseline, single-case design. During the baseline period, the intervention period, and the post-intervention period, all participants completed several health outcomes twice a week on visual analogue scales.

Setting. Rheumatology, outpatient department of a specialized hospital in the Netherlands.

Participants. One man and four women (age 51 to 76) diagnosed with GOA.

Primary outcome measures. To assess feasibility we assessed the number of drop-outs and adverse events, adherence rates, and patient satisfaction.

Secondary outcome measures. To assess the potential effectiveness ~~preliminarily~~ we assessed pain and self-efficacy~~. Effectiveness was preliminarily assessed~~ using visual data inspection and randomisation tests.

Results. The intervention was feasible in terms of adverse events (none) and adherence rate, but not in terms of participant satisfaction with the intervention. Visual inspection of the data and randomisation testing demonstrated no effects on pain (p = 0.93) or self-efficacy (p = 0.85).

Conclusions. The results of the present study indicate that the proposed intervention for patients with GOA was insufficiently feasible and effective. The data obtained through this multiple-baseline study has highlighted several areas in which the therapy programme can be optimised.

Article summary

Article focus:

- To evaluate the feasibility the effectiveness of a 12-week ~~multidisciplinary~~ non-pharmacological multidisciplinary intervention in patients with generalised osteoarthritis (GOA).
- To ~~preliminarily~~ evaluate ~~the~~ the potential effectiveness of a 12-week ~~multidisciplinary~~ non-pharmacological multidisciplinary intervention in patients with GOA.

Key messages:

- To date no studies are available that evaluate non-pharmacological, multidisciplinary care in individuals with GOA.
- The intervention evaluated in the present study appeared both insufficiently feasible and effective for patients with GOA.
- Several areas in which the therapy programme could be optimised were highlighted.

Strengths and limitations of this study:

- A multiple-baseline single-case design is particularly successful in demonstrating immediate effects, whereas we studied changes in health behaviour.
- Inherent to the design of the study is lower external validity due to the small number of included participants.

Funding statement:

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors

Competing interests:

All authors declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work.

Author's contribution:

Substantial contribution to the conception and design of the study: TJH, LK, LR, AAB, RAB and CHME.

Substantial contribution to the acquisition of the data: TJH, LR

Substantial contribution to the analysis and interpretation of the data: TJH, LR, LK, CHME, AAB and CHME.

Provided intellectual content whilst drafting the article: TJH, LR, LK, CHME, AAB and CHME.

Approved the final version to be published: TJH, LR, LK, CHME, AAB and CHME.

Data sharing statement:

No additional data available.

Introduction

A growing body of evidence shows that individuals with established osteoarthritis ~~who also report joint pain comorbidities~~with multiple joint involvement - often referred to as generalised osteoarthritis (GOA) - represent a relatively large subgroup of patients [1-4]. It has been suggested that these people might be in need of more intensive treatment options than patients with single joint complaints [1,5]. To the best of our knowledge, however, there are no studies that evaluate non-pharmacological, multidisciplinary care in individuals with GOA [5], warranting the development and evaluation of such a treatment programme. Therefore, we conceptualised a non-pharmacological, multidisciplinary treatment programme following a previously-described systematic procedure [6]. The intervention was based on recommendations for the management of hip and knee osteoarthritis [7-9], and was tailored to the needs of patients with ~~joint pain comorbidities~~multiple joint involvement [1]. Due to the complex nature of multiple joint involvement in OA [1-4] and the fact that guidelines for hip and knee OA recommend multiple non-pharmacological treatment modalities, an intervention was developed by a multidisciplinary team [8].

Before evaluating such an intervention in a randomised clinical trial, a pilot study is recommended [10], since evaluations are often undermined by problems of acceptability, compliance, delivery of the intervention, recruitment and retention, and smaller-than-expected effect sizes [11]. A useful study design for pilot interventions is the multiple-baseline single-case design, as it allows researchers to test the feasibility of the intervention and to make an ~~preliminary~~ assessment of its potential effectiveness with a low number of participants [12]. In a multiple-baseline design, the intervention is introduced to subjects after randomly-assigned baseline periods of different lengths, and an effect is demonstrated if the measured outcome only changes after the intervention has been introduced [13].

The primary aim of our study was to evaluate the feasibility of a ~~complex~~ non-pharmacological, ~~multi-disciplinary~~multidisciplinary intervention in patients with GOA. Our secondary aim was to ~~preliminarily~~ assess the potential effectiveness of this intervention on pain and self-efficacy.

Methods

Participants

Men and women, 40 years or older and referred to the ~~multi-disciplinary~~multidisciplinary intervention, were eligible to participate in the present study if they had been diagnosed with GOA; i.e. experiencing complaints in three or more joint groups, having at least two objective signs that indicate OA in at least two joints, and having limitations in daily functioning (Health Assessment Questionnaire-Disability Index score (HAQ-DI) ~~[~~[14]~~Fries 1980]~~ > 0.5) ~~according to the definition proposed by Hoogeboom *et al.*~~ [15]. Individuals were excluded from participation in the intervention if: 1) they were awaiting joint replacement surgery, 2) they had already participated unsuccessfully in a self-management programme for their GOA complaints, 3) their therapists suspected that they were suffering high levels of distress, 4) they did not master the Dutch language, or 5) they were illiterate. Recruitment and treatment of patients took place at the rheumatology outpatient department at the Maartenskliniek Woerden (the Netherlands).

The study protocol was reviewed and approved by the Institutional Review Board of the University Medical Centre Nijmegen (protocol number 2009/173), and did not fall within the remit of the Medical Research Involving Human Subjects Act.

Design

A randomised concurrent multiple-baseline single-case design was applied [13]. Participants completed repeated measurements during a baseline phase (phase A), a therapy-phase (phase B, 12 weeks) and a post-therapy phase (phase A'). Phase A acted as a control and was therefore compared with phases B and A'. By applying multiple baselines of varying length, observed effects of the treatment can be distinguished from effects due to chance [12,16,17],

thus increasing internal validity. The total duration of phase A and A' was set at 7 weeks for each participant, and consequently participants with a longer phase A had a shorter phase A'. Participants were randomly assigned to a baseline and post-therapy period of either 2 and 5 weeks, 2.5 and 4.5 weeks, 3 and 4 weeks, …, or 5 and 2 weeks, respectively, using the Wampold-Worsham method [18] to increase statistical power. During the total study period of 19 weeks, participants completed diary measures twice a week, resulting in a total of 38 measurement points (14 during phase A and A' and 24 during phase B). Each diary measure comprised 14 VAS scales.

Measurements

Feasibility of the intervention

To evaluate the feasibility of the intervention, we assessed: 1) number of, and reasons for, drop-out during the intervention; 2) adherence to the intervention (number of no shows); 3) occurrence of adverse events related to the intervention; 4) participants' satisfaction with the intervention (straightforward question ranging from 0 (totally dissatisfied) to 10 (totally satisfied)); and 5) participants' satisfaction with the assessment procedure (straightforward yes/no questions).

Diary Measures

Diary measures comprised 14 VAS scales (scoring range from 0 to 10). Pain and fatigue were measured by single straightforward questions. Furthermore, 12 items derived from validated questionnaires were scored on a VAS scale. Kinesiophobia was measured with four VAS scales [19]. Self-efficacy was assessed using two questions from the Arthritis Self-Efficacy Scale [20]. Acceptance of the disease was measured with two questions from the subscale Acceptance of the Illness Cognition questionnaire [21], and illness perceptions were evaluated

by two questions from the Illness Perception questionnaire [22]. To assess the specific complaints of each participant, we used the Patient-Specific Complaints questionnaire (PSK)) [23]. The most important complaint was assessed through the diary measure. For all scales, a higher score represented unfavourable outcomes. Pain and self-efficacy were our ~~primary~~ main secondary outcome measures.

Pre- and post-intervention measures

At baseline, we collected data on age, sex, level of education (low (no or primary education), medium (secondary school and/or preparatory middle-level vocational education), high (university of applied sciences and/or university)) and duration of symptoms. Prior to the start of the programme, we also assessed participant's expectations about its effectiveness on a scale from 0 to 10 (0 representing 'No expectations whatsoever"). Pre- and post-intervention measures consisted of a set of validated questionnaires. We measured fatigue with the "Subjective Fatigue" subscale of the Checklist Individual Strength (CIS) [24], on which higher scores represent greater fatigue. Self-efficacy was evaluated with the General Self-Efficacy Scale [25], where higher scores represent higher levels of self-efficacy. Acceptance and helplessness were measured using the Illness Cognitions Questionnaire (ICQ) [26], where higher scores reflect higher levels of agreement with that generic illness cognition. As no specific questionnaires are available to assess the self-reported functional status of individuals with GOA, we used generic questionnaires for both the lower and upper extremities, namely the Lower Extremity Functional Scale (LEFS) [27] and the Disability of Arm, Shoulder and Hand (DASH), respectively [28]. Higher scores on the LEFS and DASH represent lower and greater disability, respectively.

Intervention

The group-based intervention (8 persons per group) lasted 12 weeks, comprised 10 sessions of approximately 1.5 hours per session, and was provided by an occupational therapist and physical therapist. To ensure group learning the treatment program was decided to be delivered in a group setting,~~A group-based intervention was chosen to allow patient interaction.~~ The intervention aimed to increase the participants' knowledge of the disease, to optimise the participants' current lifestyle, and to enhance the participants' self-efficacy in controlling the disease.

To do so, patients received information on activity pacing, medication use, physical activity and weight reduction. Consequently, based on the received information participants set personal goals regarding ~~All participants received information on the disease and how to manage the disease (i.e., recommendations on activity pacing, medication use, physical activity and weight reduction~~all these health areas. By setting these personal goals, participants transferred the health information into practical and personally relevant therapy goals. Goal setting and monitoring was done according to ~~To enhance the participants' self-efficacy,~~ the 5-As model of behaviour change counselling ~~was used~~ [29]; a generally accepted method to enhance self-efficacy in health care settings. During each session, after the initial information session, the individual goals were monitored and discussed. To allow for positive feedback regarding the personal goals, all goals had to be achievable in brief amounts of time. Some examples of personal therapy goals were: 1. For the next three days, while at work, plan and perform 15 minutes of physical activity spread over three different time points (component Physical Activity); 2. For the next week, whilst cleaning the house, alternate (maximum of 10 minutes) between vacuum cleaning, other household chores, and rest moments (component Activity Pacing); 3. For the next week, use your pain medication (two tablets of Paracetamol (500 mg)) four times a day and monitor your pain during this period

(component Medication Use); and 4. For the next week, eat at least three days two slices of whole wheat bread as breakfast (component Weight Reduction).

~~Moreover~~In addition~~,~~ , daily activities (such as walking, sitting, standing, stair climbing and getting in and out of bed) were included in the therapeutic activity programme. Participants received information and practised ~~were enrolled in a therapeutic activity programme to~~how to perform these daily activities without overexerting the joints and muscles. ~~improve the quality of movement.~~ Participants were instructed and encouraged to implement these techniques and methods of performing the activities in their daily practice.

Finally, participants were familiarised with different kinds of sports, tailored to the participants' complaints to prevent overexertion (i.e. tai chi, brisk walking, and therapeutic fitness). An overview of the intervention is depicted in Box 1. Participants were advised to implement these recommendations in their home situation.

Data analysis

All data were entered into the data-entry program Epidata [30]. Ten per cent of the data was entered twice to establish the quality of data entry. Missing data were described.

Diary data were analysed using the 2-Standardised Deviation (SD) band method [17] (visual inspection) and randomisation tests [31]. The 2-SD band was calculated from the baseline data and graphed from the baseline phase through the intervention phase. If two or more successive data points in the intervention or post-intervention phase fell outside the bandwidth of 2 SDs, the result was considered significant [17]. As serial dependence - the extent to which scores at one point in a series are predictive of scores at another point in the same data set - can bias the visual inspection [17], we checked our data in each phase for serial dependence using the lag-1 method [12]. If data were found to be significantly correlated, we transformed the data using a moving-average transformation, in which the preceding and

succeeding measurements were taken into account [12,16]. In addition, randomisation tests for multiple-baseline single-case designs were carried out. We expected phase B and A' to be superior to phase A in terms of our health outcome assessment. Therefore our we tested the null hypothesis - that there would be no differential effect for any of the measurement times - using a randomisation test of the differences in the means between the pre-intervention phase and the intervention or post-intervention phase [17]. A p-value < 0.05 was considered statistically significant. For the pre- and post-measurements, we considered change scores of 20% on validated questionnaires as clinically relevant [32]. We used Stata/IC 10.1 for Windows for the descriptive and visual analysis of the data and R version 2.14.1 for the randomisation tests [31].

Results

Nine people were screened to participate in the study; two patients were excluded as they did not report functional disabilities (HAQ-DI < 0.5) and two patients who were eligible were unable to attend the program. Eventually, ~~F~~fi~~i~~ve participants gave written informed consent to participate in the study. One patient dropped out of the study within two weeks after the start of the study, reporting that filling out the questionnaires was too demanding for her on an emotional level. However, she did continue with the ~~multi disciplinary~~multidisciplinary intervention. The four remaining participants completed all 38 diary measures, resulting in 2,128 completed items. Six items (0.3%) were missing. Data entry errors were negligible (<0.1%). Table 1 presents the characteristics of the participants.

Feasibility of the intervention

Prior to the intervention, participants' expectations regarding the effectiveness of the intervention ranged from 5 to 7 (median = 7). Participant 3 missed three of the 10 sessions; participants 2 and 4 both missed one session. Participant 1 reported an increase in pain levels, which she ascribed to the intervention. Satisfaction with the intervention was assigned a score of 8 points out of 10 by participants 1, 2 and 4, and 7 points out of 10 by participant 3. Perceived therapy effects were assigned a score of 7, 3, 5, and 7 out of 10 by participants 1, 2, 3 and 4, respectively. All participants believed the questionnaires used in this study properly evaluated their most important issues. The remarks most frequently made by participants regarding the intervention were: 1) there were too many sessions and these were too short/brief; 2) too much verbal information; 3) too much time between two sessions; 4) too little information on acceptance of the disease; and 5) too little individualisation in the exercise sessions, and in setting and monitoring therapy goals.

Diary measures

Our primary effectiveness outcome measures were pain and self-efficacy. In the pain data, participant 3's intervention phase showed serial dependence, and that of participants 1 and 4 showed large fluctuations. Thus, we transformed these data prior to completion of visual data analysis. The 2-SD band method showed that participants 1, 2 and 4 each experienced significant deterioration in their pain scores between baseline, intervention and post-intervention phases. Participant 3 demonstrated significant improvement during the intervention phase (Figure 1), though this did not persist during the post-intervention phase. For all four participants, randomisation tests demonstrated no significant changes in pain between the pre-intervention phase and the intervention/post-intervention phase (p=0.93). Serial dependence was found in participant 4's self-efficacy data, and these data were transformed prior to the analyses. The 2-SD band method demonstrated that participant 4 experienced significantly higher levels of self-efficacy in both the intervention and post-intervention phase compared to the baseline phase. No differences were found for participants 1, 2 and 3. Randomisation testing demonstrated no statistically significant difference between the phase prior to the intervention and the phases during and after the intervention (p=0.85). Outcomes of the ~~R~~randomisation tests for our secondary effectiveness outcome measures ~~are~~were: fatigue (p=0.79), patient specific complaints (p=0.64), kinesiophobia (p=0.02), illness cognitions (p=0.69) and illness perception (p=0.60)~~ shown in Table 2~~.

Pre- and post-measurements

Table ~~3~~ 2 depicts the clinically relevant changes from baseline for each of the four participants. None of the participants reported improvement in self-efficacy. Participant 1 experienced clinically relevant deterioration in self-efficacy, upper body function and

kinesiophobia. Participant 4 reported improvements in fatigue levels, upper body function, kinesiophobia and acceptance. Both participants 2 and 3 remained stable.

Discussion

Our data suggest that the tailored, 12-week non-pharmacological, multidisciplinary intervention for patients with GOA was feasible in terms of adverse events, number of drop-outs and participation rate. On the other hand, the participants raised several critical points concerning the structure, content, and perceived benefits of the intervention. The latter was confirmed by visual inspection of the data and randomisation testing, as the intervention did not demonstrate clear-cut effects on health-related factors. Therefore, we believe the content and structure of the current intervention does not warrant further evaluation in a randomised clinical trial.

In view of the participants' remarks, we believe that the intervention should be more individually tailored. One of the remarks was that the therapeutic movement programme was not sufficiently individualised to address the participants' health problems. In a future non-pharmacological, multidisciplinary intervention, it might be of value to incorporate the results of the Patient-Specific Complaints instrument [23] in the therapeutic activity programme. Moreover, it was suggested that setting and achieving goals should be monitored more closely. To do so, participants should draw up action plans by completing goal-setting forms to formulate short-term goals, whilst being aware of potential limiting factors. In this way, personal goals could be monitored, discussed and adjusted, which in turn might increase the involvement and self-efficacy of the participants [17]. Finally, participants had relatively low treatment expectations regarding the intervention (highest score was 7 out of 10), implying that participants might have lacked an active role prior to the start of intervention. Motivation is considered one of the most important factors for the success of a self-management programme [33,34]. Therefore, to increase the effectiveness of a non-pharmacological, multidisciplinary intervention in patients with GOA, attention should be paid to participants'

motivation prior to inclusion. Furthermore, therapists could be trained in motivating and goal-setting techniques, for example motivational interviewing.

Several limitations should be taken into account when interpreting our data. First, we used a concurrent multiple-baseline single-case design to evaluate the intervention's ~~preliminary~~ potential effectiveness. This design is particularly successful in demonstrating immediate effects [35]. Since our intervention aimed to improve self-management in individuals with osteoarthritis, which is often considered challenging and time-consuming [9], our choice of study design might not be optimal, given the short evaluation period and the considerable length of the treatment programme. A second limitation was that all participants were in the same therapy group, possibly resulting in a negative group effect compromising any therapy effects. On the other hand, the traditional approach to multiple-baseline studies is for all participants to undergo treatment simultaneously [13]. This strategy is recommended as it improves internal validity, particularly in terms of history effects [36]. A third limitation, inherent to the design of the study, is that the study has lower external validity than randomised clinical trials, for which participants are usually selected to form a generalizable sample [37]. A ~~final~~ fourth limitation of this study was its inability to test the feasibility of study logistics for a randomised clinical trial (for example, recruitment rate, drop-out rate, and issues concerning randomisation) [38]. A final limitation was that we selected patients based on their medical diagnosis and functional status rather than on their scores on our main secondary outcomes (i.e. pain and/or self-efficacy). Future studies should include clinically relevant thresholds for their outcome measures in the in- and exclusion criteria.

As far as we know, we are the first to study a multidisciplinary intervention to improve self-management in people with GOA. Due to differences in study populations, our results cannot be compared with those of another study into the effect of a non-pharmacological, multidisciplinary intervention in patients with GOA after major joint replacement surgery

[39]. It is remarkable that so little research is available given the relatively high prevalence of ~~joint pain comorbidity in~~ individuals with established osteoarthritis with multiple joint involvement and its association with compromised health status [1,2].

Some consider single-case experimental designs as viable alternatives to large-scale randomised clinical trials [40,41], whereas others state the opposite [37,42]. Whilst using this design, we faced several (practical) constraints that potential users should be aware of. As yet, there is a plethora of analytical techniques for single-case data [31], with little or no consensus on the optimal way to analyse the data. In our study, we demonstrated a significant effect of our intervention on kinesiophobia using a randomisation test, whereas visual inspection showed only clear effects in one participant. Another practical consideration is that the design requires a substantial contribution from the participants. In the present study, one of the participants dropped out as she experienced additional psychological burden due to recurring questionnaires. It remains to be elucidated whether frequent assessment of health status as in the current study negatively, or perhaps positively, influences health outcomes. In our opinion, the multiple-baseline single-case study is a useful and valid alternative to the randomised pilot study, as it gives insight into the feasibility of the intervention and allows to ~~preliminary~~evaluate the intervention's potential effectiveness, allowing one to tailor the content and context of the intervention prior to conducting a randomised clinical trial. However, ~~it~~ single-case studies should only be considered an alternative to a full-sized randomised clinical trial in rare diseases or in situations where a randomised clinical trial is unfeasible or unethical, ~~due to~~because of the designs' limitations, including low external validity of the findings and the inability to correct for confounders (such as medication use, age, disease duration etc.).

An interesting finding was the marked variability in VAS scores within participants on specific outcomes. For example, three participants reported fluctuations in pain scores of

more than 4 points within a period of half a week (i.e., between two measurement points). Fluctuations in pain between two measurement points ranged from 0 to 7 points, frequently exceeding the thresholds for clinically relevant differences [43]. Such fluctuations indicate that pain in OA is far less stable than often believed and should perhaps be assessed far more frequently. As such variations are also likely to occur in randomised clinical trials, researchers should consider assessing post-intervention health outcomes at repeated time points. These outcomes could then be averaged to obtain a more stable post-intervention point estimate.

In conclusion, health providers and researchers should be aware of the lack of studies on the effectiveness of non-pharmacological and/or multidisciplinary interventions for patients with GOA. In our study, although we systematically conceptualised our intervention according to the latest evidence [7-9] and in collaboration with several health care providers, both feasibility and effectiveness of the care programme are doubtful. Therefore, the ~~current~~ therapy programme as described in this paper does not warrant evaluation in a large randomised clinical trial. ~~, although~~Since the data obtained in this multiple baseline study have highlighted several ways in which the therapy program could be optimized/improved, these changes should be implemented prior to conducting an RCT to further examine the interventions' effectiveness.

Reference List

1 Hoogeboom TJ, den Broeder AA, Swierstra BA, et al. Joint-pain comorbidity, health status, and medication use in hip and knee osteoarthritis: A cross-sectional study. Arthritis Care Res (Hoboken) 2012;**64**:54-8.

2 Suri P, Morgenroth DC, Kwoh CK, et al. Low back pain and other musculoskeletal pain comorbidities in individuals with symptomatic osteoarthritis of the knee: data from the osteoarthritis initiative. Arthritis Care Res (Hoboken) 2010;**62**:1715-23.

3 Forestier R, Francon A, Briole V, et al. Prevalence of generalized osteoarthritis in a population with knee osteoarthritis. Joint Bone Spine 2011;**78**:275-8.

4 Gunther KP, Sturmer T, Sauerland S, et al. Prevalence of generalised osteoarthritis in patients with advanced hip and knee osteoarthritis: the Ulm Osteoarthritis Study. Ann Rheum Dis 1998;**57**:717-23.

5 Conaghan P, Birrell F, Burke M, et al. Osteoarthritis: national clinical guideline for care and management in adults. London: Royal College of Physicians 2008.

6 Stukstette M, Hoogeboom T, de RR, et al. A multidisciplinary and multidimensional intervention for patients with hand osteoarthritis. Clin Rehabil 2012;**26**:99-110.

7 Zhang W, Moskowitz RW, Nuki G, et al. OARSI recommendations for the management of hip and knee osteoarthritis, part I: critical appraisal of existing treatment guidelines and systematic review of current research evidence. Osteoarthritis Cartilage 2007;**15**:981-1000.

8 Zhang W, Moskowitz RW, Nuki G, et al. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. Osteoarthritis Cartilage 2008;**16**:137-62.

9 Zhang W, Nuki G, Moskowitz RW, et al. OARSI recommendations for the management of hip and knee osteoarthritis: part III: Changes in evidence following systematic cumulative update of research published through January 2009. Osteoarthritis Cartilage 2010;**18**:476-99.

10 Craig P, Dieppe P, Macintyre S, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008;**337**:a1655.

11 Eldridge SM, Ashby D, Feder GS, et al. Lessons for cluster randomized trials in the twenty-first century: a systematic review of trials in primary care. Clin Trials 2004;**1**:80-90.

12 Backman CL, Harris SR, Chisholm JA, et al. Single-subject research in rehabilitation: a review of studies using AB, withdrawal, multiple baseline, and alternating treatments designs. Arch Phys Med Rehabil 1997;**78**:1145-53.

13 Christ T. Experimental control and threats to internal validity of concurrent and nonconcurrent multiple baseline designs. Psychology in the Schools 2007;**44**:451-9.

14 Fries JF, Spitz P, Kraines RG, et al. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;**23**:137-45.

15 Hoogeboom TJ, Stukstette MJ, de Bie RA, et al. Non-pharmacological care for patients with generalized osteoarthritis: design of a randomized clinical trial. BMC Musculoskelet Disord 2010;**11**:142.

16 Nourbakhsh MR, Ottenbacher KJ. The statistical analysis of single-subject data: a comparative examination. Phys Ther 1994;**74**:768-76.

17 Holtgrefe K, McCloy C, Rome L. Changes associated with a quota-based approach on a walking program for individuals with fibromyalgia. J Orthop Sports Phys Ther 2007;**37**:717-24.

18 Ferron J, Sentovich C. Statistical Power of Randomization Tests Used With Multiple-Baseline Designs. The Journal of Experimental Education 2002;**70**:165-78.

19 de Jong JR, Vlaeyen JW, de Gelder JM, et al. Pain-related fear, perceived harmfulness of activities, and functional limitations in complex regional pain syndrome type I. J Pain 2011;**12**:1209-18.

20 Lorig K, Holman H. Arthritis Self-Efficacy Scales measure self-efficacy. Arthritis Care Res 1998;**11**:155-7.

21 Evers AW, Kraaimaat FW, van LW, et al. Beyond unfavorable thinking: the illness cognition questionnaire for chronic diseases. J Consult Clin Psychol 2001;**69**:1026-36.

22 Bijsterbosch J, Scharloo M, Visser AW, et al. Illness perceptions in patients with osteoarthritis: change over time and association with disability. Arthritis Rheum 2009;**61**:1054-61.

23 Beurskens AJ, de Vet HC, Koke AJ, et al. A patient-specific approach for measuring functional status in low back pain. J Manipulative Physiol Ther 1999;**22**:144-8.

24 Beurskens AJ, Bultmann U, Kant I, et al. Fatigue among working people: validity of a questionnaire measure. Occup Environ Med 2000;**57**:353-7.

25 Luszczynska A, Scholz U, Schwarzer R. The general self-efficacy scale: multicultural validation studies. J Psychol 2005;**139**:439-57.

26 Maas M, Taal E, van der Linden S, et al. A review of instruments to assess illness representations in patients with rheumatic diseases. Ann Rheum Dis 2009;**68**:305-9.

27 Hoogeboom TJ. Translation and validation of the Dutch Lower Extremity Functional Scale. Submitted 2012.

28 Beaton DE, Katz JN, Fossel AH, et al. Measuring the whole or the parts? Validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. J Hand Ther 2001;**14**:128-46.

29 Glasgow RE, Emont S, Miller DC. Assessing delivery of the five 'As' for patient-centered counseling. Health Promot Int 2006;**21**:245-55.

30 EpiData Data Entry, Data Management and basic Statistical Analysis System. Odense Denmark, EpidData Association, 2000-2008. [computer program]. 2012.

31 Bulte I, Onghena P. Randomization tests for multiple-baseline designs: an extension of the SCRT-R package. Behav Res Methods 2009;**41**:477-85.

32 Escobar A, Gonzalez M, Quintana JM, et al. Patient acceptable symptom state and OMERACT-OARSI set of responder criteria in joint replacement. Identification of cut-off values. Osteoarthritis Cartilage 2012;**20**:87-92.

33 DeWalt DA, Davis TC, Wallace AS, et al. Goal setting in diabetes self-management: taking the baby steps to success. Patient Educ Couns 2009;**77**:218-23.

34 Teixeira PJ, Going SB, Houtkooper LB, et al. Exercise motivation, eating, and body image variables as predictors of weight control. Med Sci Sports Exerc 2006;**38**:179-88.

35 Neuman SB. Single-Subject Experimental Design. In: Duke NK, Mallette MH, editors. Literacy Research Methodologies. 2nd edition ed. New York: The Guilford Press; 2011. p. 383.

36 Carr JE. Recommendations for reporting multiple-baseline designs across participants. Behav Intervent 2005;**20**:219-24.

37 Newcombe RG. Should the single subject design be regarded as a valid alternative to the randomised controlled trial? Postgrad Med J 2005;**81**:546-7.

38 Thabane L, Ma J, Chu R, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol 2010;**10**:1.

39 Bulthuis Y, Drossaers-Bakker KW, Taal E, et al. Arthritis patients show long-term benefits from 3 weeks intensive exercise training directly following hospital discharge. Rheumatology (Oxford) 2007;**46**:1712-7.

40 Rizvi SL, Nock MK. Single-case experimental designs for the evaluation of treatments for self-injurious and suicidal behaviors. Suicide Life Threat Behav 2008;**38**:498-510.

41 Janosky JE. Use of the single subject design for practice based primary care research. Postgrad Med J 2005;**81**:549-51.

42 Guyatt GH, Keller JL, Jaeschke R, et al. The n-of-1 randomized controlled trial: clinical usefulness. Our three-year experience. Ann Intern Med 1990;**112**:293-9.

43 Lee JS, Hobden E, Stiell IG, et al. Clinically important change in the visual analog scale after adequate pain control. Acad Emerg Med 2003;**10**:1128-30.

Box 1. Pat-plot of the ~~multi-disciplinary~~multidisciplinary intervention.

Timeline		Intervention
Pre-measurement		
Week 1	Part 1	a
	Part 2	b c
Week 2	Part 1	d e
	Part 2	b c
Week 3	Part 1	d 1
	Part 2	c 1 2
Week 4		e 1 2
Week 5		b d 1
Week 6		c 1 2
Week 7		e 1
Week 9		b c 1
Week 12		d 1 f
Post-measurement		

a	Introduction meeting.
b	Information on Pain and Medication use.
c	Information on Activity Pacing.
d	Information on the importance of Physical Activity.
e	Information on Weight Reduction.
1	Activity programme to improve quality of movement.
2	Sports activity.
f	Evaluation time point.

1

Table 1. Characteristics of the study participants.

Participant	Sex	Age (y)	Education	No. painful joint groups (0 - 11)	Baseline assignment (measurements)
1	F	76	Low	8	4
2	F	68	Medium	3	5
3	M	59	Low	11	7
4	F	56	High	5	6
5[†]	F	51	High	-	6

Abbreviations: F, female; M, male; No., number of.

[†] Dropped out.

2

3

1

Table 2. Clinically relevant differences between baseline and post-intervention measurements.

	Fatigue		Self-efficacy		Function				Kinesiophobia		Illness Cognitions			
					Upper		Lower				Help		Accept	
	T0	T1	T0	T1	T0	T1	T0	T1	T0	T1	T0	T1	T0	T1
pt1	42	39	35	27	35	50	44	47	43	50	11	11	12	12
pt2	9	9	35	37	**18**	**13**	69	68	28	31	8	9	23	24
pt3	**56**	**33**	35	30	31	43	38	41	57	53	13	14	**15**	**19**
pt4	**34**	**27**	29	31	**44**	**32**	46	48	**48**	**34**	9	9	**11**	**14**

Bold = 20% improvement, Underlined = 20% deterioration. Abbreviations: Accept = Subscale Acceptance; Help = Subscale Helplessness; Lower = Lower extremity functioning; pt# = Participant #; T0 = Baseline measurement; T1 = Post-intervention measurement; Upper = Upper extremity functioning.

2

3

4

Figure 1. Diary measures for pain with 2-SD horizontal band graph for baseline (phase A), intervention (phase B) and post-intervention (phase A') phases. Scores on the pain VAS range from 0 to 10; higher scores indicate higher levels of pain.

Figure 2. Diary measures for Self-Efficacy with 2-SD horizontal band graph for baseline (phase A), intervention (phase B) and post-intervention (phase A') phases. Scores on the pain VAS range from 0 to 10, higher scores indicating lower levels of self-efficacy.

We like to thank both Nadine Foster and Bhasker Amatya for their remarks and comments on our manuscript. We strongly believe the article has improved considerably regarding its quality, clarity and reproducibility and that we were able to incorporate the suggestions successfully.
The following list shows in detail how we dealt with each of the problems that the reviewers noted.We want to point out to the reviewers that the references made to page and line numbers comply with the marked manuscript.

Reviewer 1: Nadine Foster

Reviewer 1's remark 1:
Why was a multidisciplinary intervention selected as the intervention - there is no clear justification for this given in the paper. Was this intervention already available or was it developed specifically for this research study?

Comment to reviewer 1's remark 1:
We have added our justification for selecting a multidisciplinary intervention to the article as well as the statement that the intervention was specifically developed for this study (please see page 5, lines 13-16).

Reviewer 1's remark 2:
Similarly why was the intervention group-based? No justification is given for this.

Comment to reviewer 1's remark 2:
We have added a justification on why the intervention was group-based to the manuscript, please see page 10, line 3-4.

Reviewer 1's remark 3:
The intervention summary box is useful but highlights that it really mostly comprised information-giving/education. Yet we know from previous research in clinical conditions that education is a rather weak intervention to change behaviour. Thus the authors need to justify the components of the intervention more clearly. Also specifically what was the activity programme - did it focus on best evidence to date in focusing on strengthening and aerobic exercise? The papers says 'focus on quality of movement' - what is meant by this and why was this the focus rather than strengthening exercise (for which there is most evidence for effectiveness in OA)? Also the authors state the intervention was tailored but do not provide any information on how it was tailored? Would some specific examples be useful. It must be challenging to truly tailor a group-based intervention?

Comment to reviewer 1's remark 3:
We initially kept this section brief due to fact that we included the Pat-plot in our manuscript, however we agree with the reviewer's remarks that some of the aspects are too briefly described and need further clarification. Therefore we rewrote most of the 'Intervention'-paragraph, and added information addressing the reviewer's concerns. Please see paragraph Intervention (page 10-11).

Reviewer 1's remark 4:
The exclusion criteria 'if therapists suspected high levels of distress' is unclear and unjustified. How was it assessed? What is meant by it?

Comment to reviewer 1's remark 4:

If the therapists believed the patients with high distress levels would negatively impact the group process, patients were excluded from the group-based programme and offered an individual intervention. This clarification was included in the manuscript (page 7, line 12). Since the additional value of the use of validated questionnaires as a screening instrument for this purpose has not yet been proven, these were not incorporated in the present study and judgments were based on clinical experience.

Reviewer 1's remark 5:
Page 7 states that adherence was measured but the paper never explains how.

Comment to reviewer 1's remark 5:
Adherence to the multi-disciplinary therapy was determined by determining the number of no-shows to the actual therapy. We added this to page 8, line 13.

Reviewer 1's remark 6:
Page 10 - why was 20% change deemed clinically relevant? Whilst it seems reasonable, other research has shown a need for 30% or more. Again, what is the justification for 20%?

Comment to reviewer 1's remark 6:
We wanted to comply with the cut-off's used by the OARSI-responders criteria [Escobar A, Gonzalez M, Quintana JM, Vrotsou K, Bilbao A, Herrera-Espiñeira C, Garcia-Perez L, Aizpuru F, Sarasqueta C. Patient acceptable symptom state and OMERACT-OARSI set of responder criteria in joint replacement. Identification of cut-off values. Osteoarthritis Cartilage. 2012 Feb;20(2):87-92. Epub 2011 Nov 20.]. We added the reference to the manuscript (page 12, line 8).

Reviewer 1's remark 7:
The team selected a feasibility study or different design to the future hoped for RCT. Why was a pilot RCT not carried out if the ultimate plan was to inform a main RCT?

Comment to reviewer 1's remark 7:
As discussed in the manuscript, both the pilot RCT and single case study provide useful data for preparing a large RCT regarding the feasibility of the intervention as well as preliminary information on its effectiveness. As we were more interested in the feasibility of the intervention, rather than for example issues with randomization or sampling we decided to choose the design of the single-case study.

Reviewer 1's remark 8:
Overall the sample size, even for single case research, is small (only 4 of 5 provided data) and ultimately the study is based on only one small group that received the intervention as a group of OA patients.

Comment to reviewer 1's remark 8:
We agree with the reviewer that the sample is fairly small and have discussed this throughout the paper (see for example the Article Summary - Strengths and limitations of this study). We believe, however, that despite the small sample size, the present study provides useful information on the intervention and points for improvement. Furthermore, the study underlines the importance of piloting interventions and therefore serves as an example for other researchers.

Reviewer 1's remark 9:

I didn't quite follow the authors argument that the research shows they should not do a main RCT, I would have thought that the research shows clearly that the content and process of delivery of the intervention needs significant re-thinking but that ultimately a future main RCT would still be the right way to move forward to test its effectiveness.

Comment to reviewer 1's remark 9:
We agree with the reviewer that, although points of improvement were found for the present intervention, a RCT should ultimately be conducted to further study the effectiveness of multi-disciplinary interventions for GOA. What we meant with our conclusion was, that the intervention as described in this paper should not be evaluated in a randomized clinical trial, as it will most likely result in disappointing outcomes and there is room for improvements. To make this clearer, we have adjusted our manuscript's conclusion. Please see page 19, line 16-17.

Reviewer 1's remark 10:
Reference 6 is missing some details

Comment to reviewer 1's remark 10:
Thank you, the paper has just now been published and can be referred to in more detail.

Reviewer 1's remark 11:
Reference 14 refers to a RCT protocol - I was confused by this. Is this protocol for a different RCT with a different intervention?

Comment to reviewer 1's remark 11:
This reference describes the protocol for a RCT, in which a different multidisciplinary intervention is tested than described in this paper.

Reviewer 1's remark 12:
Table 1 - how was education level determined?

Comment to reviewer 1's remark 12:
We have added the meaning of the education levels Low, Medium and High education to the text (page 9, line 8-10).

Reviewer 1's remark 13:
Table 2 seems a bit meaningless with only *p*-values; could average data summary statistics be added?

Comment to reviewer 1's remark 13:
We agree with the reviewer that Table 2 seemed a bit meaningless the way it was presented in the manuscript. However, we do not think adding average data summary statistics would be a solution, as these data (n=4) will add very little information. Therefore, we decided to remove this table and implement its content in the manuscript's text (please see page 14, lines 18-20).

Reviewer 1's remark 14:
Table 3 needs a fuller footnote explaining all abbreviations

Comment to reviewer 1's remark 14:
We have clarified Table 3 by expanding the footnote. (Note: Table 3 is now Table 2)

Reviewer 1's remark 15:
Figures - label phases a, b and A'

Comment to reviewer 1's remark 15:
We have updated our figures (and their legends) according to the reviewer's recommendation.

Reviewer 2: Bhasker Amatya

Reviewer 2's remark 1:
Research question is vague and confusing, needs to be shortened and indication of patient population is missing. For example, the term "preliminary effectiveness" is not explicable. I would suggest the review of the title for e.g. 'Feasibility and effectiveness of a non-pharmacological MD care programme for persons with GOA: a randomised multiple-baseline single-case study".

Comment to reviewer 2's remark 1:
We have adjusted the title along the recommendations of the reviewer. In addition, we removed the word "preliminary" from our manuscript and replaced it by the term "potential".

Reviewer 2's remark 2a:
Not sure if this is the appropriate design, as main aim of the study as anticipated by the authors are feasibility and effectiveness of the MD programme. I am not sure how feasibility can be assessed using this design, as measuring the dependent variable prior to administering treatment is an important aspect of this type of study.

Comment to reviewer 2's remark 2a:
We understand the reviewer's concerns, but we do not fully agree with them. We believe the single-case design can be used to investigate the feasibility of an intervention, as long as the limitations of the single-case design are taken into account. For example, this design does not allow researchers to test issues regarding randomization or to determine the number of eligible non-volunteers (we have discussed this in the paper). On the other hand, it does allow to study the feasibility of the intervention itself and to determine whether evaluation of the program in a large randomized clinical trial would be worthwhile, or that further adjustments to its content are warranted.

Reviewer 2's remark 2b:
The effectiveness can sure be measured to some extent, however, the authors did not explain how the severity of the problem is quantified with measurement of the pain in a baseline period before treatment is introduced (as it seems pain scores in a VAS scale seems to be low threshold at baseline in majority of patients- in 3 out of 4).

Comment to reviewer 2's remark 2b:
In this study patients had to report functional disabilities in their daily living (HAQ-DI score of 0.5 or higher); this is part of the GOA definition which is now added to the manuscript. However, we do agree with the reviewer that additional thresholds for pain and/or self-efficacy levels would have been of value in selecting patients eligible for the intervention. We therefore have addressed this point of concern in our limitations paragraph in the discussion (page 17, lines 18-21).

Reviewer 2's remark 2c:
In addition, the A-B-A design assumes that when treatment is withdrawn, the condition would return to at least nearly what it was before the treatment began. However, with the multidisciplinary interventions the authors suggested usually we would expect to have a more lasting effect for longer-time, requiring longer follow-up. Furthermore, confounding variables (medication, age, disease duration etc.) is usually not possible with this design, and there is possibility that these confounding factors other than the treatment could have influenced the result.

Comment to reviewer 2's remark 2c:
Indeed, the A-B-A' design assumes that the therapy effect should vanish after the B-phase. We however specifically chose for this design-type as this allowed us to more clearly distinguish between the treatment phase (B-period) and the post-treatment phase (A'), as many single case studies describe the study effectiveness during the intervention phase (B-period). So, even though the design type might imply that we expected the therapy results to disappear, we explicitly describe that we expect the effect to be superior to the initial phase (A-period) (see page 11). We could describe the whole article as if we have used an A-B design, however that way we would have to eliminate a whole number of interesting data points.
The point raised by the reviewer that the study design does not allow researchers to correct for potentially confounding factors is true. As we find it important to point this out to the reader, we have stated this in our discussion section (page 18, line 21-23).

Reviewer 2's remark 3:
Definition of the GOA needs to be elaborated (Methods section, first paragraph: line 11-14)

Comment to reviewer 2's remark 3:
We have now stated the definition of GOA in the paper (please see page 7, lines 6-8).

Reviewer 2's remark 4:
Not consistency with the primary and secondary measures throughout the abstract and text. e.g. in abstract the authors indicates that feasibility as a primary outcome and effectiveness as secondary. However, in text in multiple occasions pain and self-efficacy are indicated as primary outcomes.

Comment to reviewer 2's remark 4:
This is the result of unclear writing. Feasibility is the primary outcome, but pain and self-efficacy are the main outcomes of interest in our research question on the effectiveness of the intervention. We have changed these vague statements throughout the manuscript (please see page 9, line 4-5 and page 14, line 3 & 18).

Reviewer 2's remark 5:
Interventions: not consistent throughout the text. Please note non-pharmacological and multidisciplinary (MD) intervention are two broad terms and have diverse definition. For e.g. non-pharmacological intervention range from exercise/physical modalities to orthotics and education, where as MD intervention might be non-pharmacological and pharmacological, as well as non-pharmacological programme only provided by more than 2 disciplines. Needs to define the intervention in more details and needs to be consistent.

Comment to reviewer 2's remark 5
We have changed these inconsistencies throughout the manuscript, now labeling our intervention as a "non-pharmacological, multidisciplinary intervention". Moreover, we have described our intervention into more detail on page 10 and 11.

Reviewer 2's remark 6:

The authors statement in key message (first dot point and in introduction): '…no-studies are available that evaluate non-pharmacological care in GOA' seems not accurate, as there are lots of systematic reviews and studies evaluating these interventions in OA, which can be generalised to the GOA.

Comment to reviewer 2's remark 6:
We agree with the reviewer that there are systematic reviews of interventions in OA, but as far as we know, none of those reviews actually provide data on persons with GOA. For this reason, the National Institute for Health and Clinical Excellence (NICE) included a statement in their OA guideline that trials in specifically people with GOA are absent and need to be performed. Therefore, we believe our statement is accurate.

Reviewer 2's remark 7:
Joint pain is the cardinal sign of any OA including GOA, not comorbidities as stated in Introduction, should this be 'generalised joint-pain' instead? Please review.

Comment to reviewer 2's remark 7:
We have changed the phrasing of joint-pain comorbidities into multiple joint involvement, which is more accurate in this context.

Reviewer 2's remark 8:
Practical (or clinical) significance of the findings is not clear as it seems the intervention has not made a meaningful difference in the well-being of the participant. However, authors comment in Discussion section stating that '…current intervention does not warrant further evaluation in RCT' is arguable. As this might be due to the study design itself as the intervention was provide in a group and not tailored to patient needs and goal oriented.

Comment to reviewer 2's remark 8:
Even though the intervention was group-based, we did tailor the different aspects of the intervention to the individual health needs by means of goal setting. We have made our intervention more clear and reproducible by describing it into more detail in the Intervention paragraph (Page 10 and 11).

Reviewer 2's remark 9:
It is well recognised that the sample consisting of a single subject engaged in a particular intervention provided by a particular individual is challenging, particularly in this study, due to the broad nature of the intervention. Usually, direct replication, systematic replication, and clinical replication is required for generalizability of the results from single-subject designs. Trialling the intervention using other study design with more participants and a control group would be ideal.

Comment to reviewer 2's remark 9:
We agree that research on therapy options in this group of patients should not be aborted due to the negative results found in this study. However the studied intervention in its current form needs some rethinking before we re-evaluate it in scientific study. We changed our conclusion (page 19, line 13-17) accordingly to make this point clear to the reader.

Reviewer 2's remark 10:
Introducing the patient recruitment procedure at the beginning might be helpful to the reader. How many were asked to participate, how many refused.

Comment to reviewer 2's remark 10:
We have added this information to the paper, please see page 13, lines 3-5.

Reviewer 2's remark 11:
Feasibility of the program is arguable as the median expectation of participant prior to the programme (md=7) and perceived therapy effects (md=6).

Comment to reviewer 2's remark 11:
We agree with the reviewer. We have discussed this in the second paragraph of our discussion and one of the key messages states this as well.

Reviewer 2's remark 12:
The authors fail to set a cut-off score for both pain and self-efficacy, which would have aid to inspect for changes in level (magnitude) or reductions in variability.

Comment to reviewer 2's remark 12:
We agree with the reviewer on this point. As stated earlier (Reviewer 2's remark 2b), we have addressed this issue in our limitation section of the discussion (please see page 17, lines 18-21).

Reviewer 2's remark 13:
Risk that evaluator bias and/or demand characteristics of the patients (e.g. not motivated) needs to be addressed as this might have influence the results.

Comment to reviewer 2's remark 13:
It is not likely that an evaluator bias occurred in our study, as most of the measures were completed at the participants' home. Also, the pre- and postintervention questionnaires were send out by mail. However the impact of patients' characteristics on the (lack of) treatment effects is indeed important. We have added this statement to the discussion section (please see page 18, lines 21-23).

Reviewer 2's remark 14:
The discussion section should include, What is the take home message for readers?

Comment to reviewer 2's remark 14:
We agree a take home message is important, however we believe the take home message is described pretty clearly in the Article Summary – Key Messages section of the paper (please see page 3).

Reviewer 2's remark 15:
Figures need modifications: needs to indicate the A-B-A' in all figures.

Comment to reviewer 2's remark 15:
We have updated our figures (and their legends) according to the reviewer's recommendation.