



Published in final edited form as:

Pain. 2008 November 15; 140(1): 145–157. doi:10.1016/j.pain.2008.07.029.

A Randomized Trial of Behavioral Physical Therapy Interventions for Acute and Sub-Acute Low Back Pain (NCT00373867)

Steven Z. George¹, Giorgio Zeppieri Jr.³, Anthony L. Cere³, Melissa R. Cere³, Michael S. Borut³, Michael J. Hodges³, Dalton M. Reed³, Carolina Valencia¹, and Michael E. Robinson²

¹Department of Physical Therapy, University of Florida, Gainesville FL

² Department of Clinical and Health Psychology, University of Florida, Gainesville FL

³Department of Physical Therapy, SHANDS Rehab, Gainesville FL

Abstract

Psychological factors consistent with fear-avoidance models are associated with the development of chronic low back pain (LBP). As a result, graded activity (GA) and graded exposure (GX) have been suggested as behavioral treatment options. This clinical trial compared the effectiveness of treatment based classification (TBC) physical therapy alone, to TBC augmented with GA or GX for patients with acute and sub-acute LBP. Our primary hypothesis was that GX would be most effective for those with elevated pain-related fear. In total, 108 patients enrolled in this clinical trial and were randomly assigned to receive TBC, GA, or GX. Outcomes were assessed by a blinded evaluator at 4-weeks and by mail at 6-months. The primary outcomes for this trial were disability and pain intensity, and secondary outcomes were fear-avoidance beliefs, pain catastrophizing, and physical impairment. There were no differences in 4-week and 6-month outcomes for reduction of disability, pain intensity, pain catastrophizing, and physical impairment. GX and TBC were associated with larger reductions in fear-avoidance beliefs at 6-months only. Six month reduction in disability was associated with reduction in pain intensity, while 6-month reduction in pain intensity was associated with reductions in fear-avoidance beliefs and pain catastrophizing. This trial suggests that supplementing TBC with GA or GX was not effective for improving important outcomes related to the development of chronic LBP.

Introduction

Psychological factors are associated with the development and maintenance of chronic low back pain (LBP)[35,36]. As a result, biopsychosocial models of LBP[48,49] have been endorsed as an effective way to manage LBP. The Fear-Avoidance Model of Musculoskeletal Pain (FAM) is a specific psychological model used to explain the development and maintenance of chronic LBP [25].

The FAM proposes that pain-related fear (including fear of movement and re-injury) and pain catastrophizing are the primary affective and cognitive components influencing pain

Correspondence Steven Z. George PT, PhD, PO Box 100154, University of Florida, Gainesville, FL 32611-0154, Phone: (352) 273-6432, Fax: (352) 273-6109, szgeorge@php.ufl.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

perception. These factors interact to determine the individual's initial behavioral response to pain, which can range from avoidance (maladaptive) to confrontation (adaptive). Long-term avoidance behavior has been hypothesized to have adverse psychological, physical, and societal consequences[46], although evidence for this hypothesis is not absolute [40].

Effective behavioral interventions based on the FAM have been reported in the literature [4, 18,28,44,47]. Graded exercise and graded exposure are specific behavioral interventions that dose exercise and activity parameters on factors other than pain abatement. Briefly, graded exercise utilizes a quota system to progress subjects' therapeutic exercise and activity [13]. In contrast, graded exposure hierarchically exposes subjects to specific situations of which they are fearful by starting with exercise or activity that elicits minimal amounts of fear, and then gradually increasing to situations that elicit larger amounts of fear [44].

Currently, there is a lack of consensus on whether graded exercise or graded exposure is a more effective intervention for patients with LBP. There is also a question of whether behavioral interventions should be targeted toward those with elevated pain-related fear [43], as such individuals may have a higher likelihood of developing chronic LBP. These are important issues to resolve if behavioral interventions are to be used as secondary prevention strategies in patients with acute or sub-acute LBP[14,22,29,31].

Therefore, the purpose of this randomized clinical trial was to compare the effectiveness of physical therapy to physical therapy augmented with graded exercise or graded exposure for patients with acute and sub-acute LBP. The primary outcomes for this clinical trial were disability and pain intensity, and the secondary outcomes were fear-avoidance beliefs, pain catastrophizing, and physical impairment. Our previous randomized trial reported that treatment based classification physical therapy augmented with graded exercise was effective only for those with elevated pain-related fear[18]. Subsequent reports have suggested that graded exposure might be more effective for patients with chronic LBP [27,30,44,45] leading to its inclusion in the current trial. Our primary hypothesis was that graded exposure would be most effective for those with elevated pain-related fear, in comparison to the other treatments. Additionally, we hypothesized that reduction in fear-avoidance beliefs and pain catastrophizing would be predictive of reduction in disability and pain intensity, lending additional support for their inclusion as treatment targets for patients with acute and sub-acute LBP[19].

Methods

Overview

Patients referred for rehabilitation of LBP that met the eligibility criteria and provided informed consent were randomly assigned to receive treatment based classification physical therapy alone, or augmented with graded exercise, or graded exposure. Licensed physical therapists followed treatment guidelines for 4-weeks. Patients were re-assessed by a blinded evaluator at 4-weeks and provided mail follow-up at 6-months after the date of random assignment. The relative effectiveness of the interventions was determined by statistically comparing the three groups for differences at 4-weeks and 6-months. This clinical trial was registered at www.clinicaltrials.gov under the identifier of NCT00373867.

Subjects

Consecutive patients seeking treatment for LBP at 3 participating University of Florida affiliated clinics were considered for participation from January 2006 to October 2007. This trial was approved by the University of Florida Institutional Review Board, and all subjects provided informed consent before study participation was confirmed. The inclusion and exclusion criteria were based on guidelines from the Quebec Task Force on Spinal Disorders

(QTFSD)[1]. For purposes of this trial acute and sub-acute LBP were operationally defined as reporting current symptoms for 1 – 24 weeks. Chronic LBP was defined as reporting current symptoms for greater than 24 weeks.

Inclusion Criteria

- QTFSD classification
 - 1a or 1b (acute or sub-acute LBP without radiation below the gluteal fold);
 - 2a or 2b (acute or sub-acute LBP with proximal radiation to the knee)
 - 3a or 3b (acute or sub-acute LBP with distal radiation below the knee)
- The ability to read and speak English
- Age 15 – 60

Exclusion Criteria—Patients were excluded for meeting any one of the exclusion criteria.

- QTFSD classification
 - 1c (chronic LBP without radiation below the gluteal fold)
 - 2c (chronic LBP with proximal radiation to the knee)
 - 3c (chronic LBP with distal radiation below the knee)
 - 4a or 4b or 4c (acute or subacute or chronic low back pain with distal radiation below the knee and neurological signs)
 - 5 (presumptive lumbar nerve root compression)
 - 6 (confirmed lumbar nerve root compression)
 - 7 (confirmed lumbar spinal stenosis)
 - 8 (post-surgical status, less than 6 months after surgical intervention)
 - 9.1 (post-surgical status, more than 6 months after surgical intervention, asymptomatic)
 - 11 (other spinal disorders including metastatic disease, visceral disease, or fracture)
- Pregnancy
- Osteoporosis

Measures

Measures were administered prior to randomization at baseline and 4-weeks after randomization. Patients completed 6-month follow-up data by mail.

Demographic and historical information—Demographic and historical data was collected with a standard questionnaire. Information collected included gender, age, employment status, marital status, educational level, onset of symptoms, duration of symptoms, and previous episodes of low back pain.

Treatment preference—Patient preference for treatments in this study was measured with a standard form. These data were collected because patient treatment preference may influence the outcome of a randomized trial by affecting response to treatment or by enrolling patients without strong treatment preferences [11,42].

Treatment satisfaction—Treatment satisfaction was assessed with selected items from a previously described measure that assessed treatment delivery and effect [7]. At the 6-month assessment, patients mailed their questionnaire responses directly to the investigators to ensure clinical staff would not influence satisfaction ratings.

Primary outcome measures—Primary outcomes were assessed with previously validated measures of low back related disability and pain intensity. Disability was assessed with the Oswestry Disability Questionnaire (ODQ), which has been recommended as appropriate outcome measure for self-report of disability[9,10]. The ODQ has 10 items that assess how LBP affects common daily activities, for example sitting, standing, and lifting. The ODQ has a range of 0 (no disability due to back pain) to 100 (completely disabled due to back pain), so higher scores indicate higher disability from LBP. Patients rated their pain intensity using a numerical rating scale (NRS), ranging from “0” (no pain) to “10” (worst pain imaginable) [23]. Patients rated pain intensity over 3 conditions, the present pain intensity, the worst pain intensity over the past 24 hours, and the best pain intensity over the past 24 hours. These 3 ratings were summed and divided by 3 (arithmetic mean) for use in data analyses. Use of this composite score is justified by its high correlation with actual average pain intensity ratings ($r = .82$) from a previous study using daily diary ratings as reference standard[24].

Secondary outcome measures—Secondary outcomes were assessed with previously validated measures. The Fear-Avoidance Beliefs Questionnaire (FABQ) was used to quantify fear-avoidance beliefs[50]. The FABQ contains 2 scales, a 7-item FABQ work scale (FABQ-W, range 0-42) and a 4-item FABQ physical activity scale (FABQ-PA, range 0-24). Higher scores indicate higher levels of fear-avoidance beliefs for both FABQ scales. The FABQ-PA was used in this trial because it is more commonly used as an outcome measure (including our prior study) and we anticipated recruiting low amounts of work related LBP from these clinical settings. The Pain Catastrophizing Scale (PCS) was used to quantify pain catastrophizing. [41] The PCS is a 13-item scale that assesses the degree of catastrophic cognitions a patient reports due to LBP. The PCS has a total range of 0 to 52 and higher scores are associated with higher amounts of pain catastrophizing. A physical impairment scale (PIS) performed by blinded physical therapist quantified physical impairment due to LBP[51]. The PIS consists of 7 different examination procedures performed by the patient and performance for each procedure is scored as a negative (0) or positive (1) for presence of impairment. The PIS has a total range of 0 to 7, and higher scores indicate higher levels of physical impairment due to LBP. The PIS was not performed at 6-months because this follow-up was performed by mail.

Randomization Procedure

The randomization scheme was prepared by computer and completed prior to the start of the study. Randomization was balanced to ensure equal allocation to each treatment group, and randomization was stratified based on the 3 participating clinics. Treatment assignments were contained in sealed, numbered envelopes which were opened in sequential order, as each patient enrolled in the study.

Physical Therapist Selection and Training

Licensed physical therapists were selected from a pool ($n = 11$) that participated in an in-service given by the first author. These therapists completed a survey about their beliefs of treatment effectiveness for the interventions being studied. A majority believed the treatments had an average or better effect; 9 (82%) for the standard care, 10 (91%) for the graded exercise, and 11 (100%) for the graded exposure. The 6 physical therapists selected for this study were matched on having average or better treatment effectiveness beliefs for all 3 interventions. Three physical therapists worked in the first clinic, 2 worked in the second clinic, and 1 worked at the third clinic. These physical therapists completed a training module led by the first author

that included instruction in standard allocation of the interventions, documentation for research purposes, and procedure for reporting adverse events. The physical therapists participating in the study demonstrated written competence in all these areas.

Intervention

Intervention details are summarized below and in Table 1. The intervention was considered allocated for a given subject if he or she attended 2 or more physical therapy sessions. The total time for each intervention session ranged between 45 minutes – 1 hour, regardless of the assignment.

Treatment based classification physical therapy—All patients received physical therapy consistent with a treatment-based classification (TBC) system originally described by Delitto et al [8]. TBC is a standard system with acceptable reliability that assists clinicians in identifying LBP sub-groups and matching appropriate treatment[15,17]. These sub-groups have been previously described in more detail and are named for their corresponding treatment category of specific exercise, manipulation or mobilization, lumbar stabilization, or traction [17]. Furthermore, these LBP subgroups were used in our previous clinical trial[18], and other recently reported clinical trials [3,16]. Patients assigned to TBC had patient education materials that were anatomically focused, and supplemented with the *Handy Hints* pamphlet[18].

Graded activity—Patients randomly assigned to this group had their TBC augmented with graded exercise (GA). The GA protocol followed parameters established in our previous clinical trial[18]. Briefly, an activity quota was determined by having the patient perform standard aerobic and therapeutic exercises to pain tolerance. The parameters (duration, intensity, and frequency) used to reach pain tolerance were then established as the activity quota. Graded activity principles were used to progress exercise during subsequent treatment sessions[13]. For example, if the patient met his/her quota it was increased by at least 10% in duration, intensity, and/or frequency for future treatment sessions. If the patient did not meet the quota it was maintained for the next session. Patients assigned to GA had patient education materials that were focused on a biopsychosocial model, and supplemented with *The Back Book* pamphlet[37].

Graded exposure—Patients randomly assigned to this group had their TBC augmented with graded exposure (GX). The GX protocol was a modified version of what has been previously described in the literature by Vlaeyen et al[44]. First, fearful activities were assessed with a standard questionnaire that listed 10 activities subjects with LBP are commonly fearful of, for example, lifting, carrying, twisting, and bending. The questionnaire also had 2 options for open-ended responses where the subject could provide additional examples. Patients rated each item using a numerical rating scale NRS that ranged from 0 (No fear) to 100 (Maximal fear). A summary of the 10 activity fear ratings is included in Table 2.

The physical therapists then selected 2 items ranked as most fearful for implementation in the GX protocol. Lifting more than 20 pounds, carrying more than 20 pounds, and reaching to floor were commonly rated as highly fearful by this sample (Table 2). Standardization of the activities was not possible because of the variability in clinical presentation and level of functional activity by a particular patient. However, the progression of activity followed GX principles and was standardized to address components related to position, intensity, frequency, and duration as appropriate for a given activity. For example, a patient that was fearful of reaching to the floor would include progression in position (lumbar flexion in supine to sitting to standing), frequency (lower to higher repetitions), and intensity (no resistance to resistance).

All GX activities were completed under supervision of study physical therapists and appropriate clinical staff. Continued interaction with the patient to obtain NRS fear ratings following activity performance was also an important part of the GX protocol. Activities rated as highly fearful were initially included in the treatment plan at positions, intensities, frequencies, and durations that did not increase fear. This initial level of activity was determined through patient history or by having the patient perform the activity at a self-selected level. The decision to increase or maintain exposure levels was made from NRS fear ratings that occurred within the same treatment session. Patient activity was maintained if NRS fear rating did not change or increased after exposure to fearful activity. In this situation the importance of attempting these activities was explained to the patient and activity exposure was maintained at the current position, intensity, frequency, and duration for the subsequent session.

Patient activity was progressed if a decreased NRS fear rating was reported following exposure to a fearful activity. In this situation the patient received positive reinforcement for confrontation. Then, the position, intensity, frequency, and/or duration of activity was increased for the subsequent session. The physical therapists increased activity levels by at least 10% as per standard guidelines for this clinical trial. Patient feedback was also obtained during this process to ensure that the increased activity level was associated with higher fear NRS ratings. Once fear levels were adequately decreased in the clinic, patients were encouraged to expose themselves to these activities as part of the home exercise program. Patients assigned to GX had patient education materials that were focused on a biopsychosocial model, and supplemented with *The Back Book* pamphlet[37].

Sample Size Estimate

Sample size estimates were based on the previous clinical trial in which a four variable regression model explained 33% of the total variance in 6-month disability outcome[18]. We estimated that a treatment interaction between the baseline level of fear-avoidance beliefs and type of treatment would need to add between 5-7% additional variance to be a significant addition to the regression model. Using those parameters, 80 subjects were necessary to detect this interaction effect with statistical power of 84%. 6-month follow up rates were estimated to be at 85%, therefore, to maintain 84% statistical power the total number of subjects necessary to recruit was 92.

Data Analysis

Analyses were performed with SPSS, version 15.0 (Chicago, IL) and reported results were consistent with CONSORT recommendations[33]. Baseline descriptive statistics for each intervention group were calculated for demographic, historical, treatment preference, primary, and secondary outcome measures. Then, post-randomization differences were assessed using ANOVA for continuous data and chi-square tests for categorical data. Variables that differed between intervention groups were considered as covariates in subsequent analyses.

Dropout and treatment satisfaction rates were assessed across the intervention groups to assess systematic differences due to treatment. We analyzed the data in 2 different ways to assess the effect missing data had on our hypotheses. First, we analyzed the data for completers only (i.e. as treated analysis), as a liberal estimate of treatment effect. Then, we used an intention to treat (ITT) principle (last value forward) to complete the analyses and generate a conservative estimate of treatment effect.

The primary hypothesis was investigated with repeated measures ANOVA within the general linear model using pre-defined main effects consistent with the previous clinical trial[18]. The main effects were initial pain related fear level (dichotomized high and low based on FABQ-PA > 14), randomly assigned treatment group (TBC, GA, or GC), and time (baseline, 4-weeks,

and 6-months). The 3-way interaction between initial fear level, treatment group, and time interaction was the effect of interest to support or refute our primary hypothesis. Separate models were created for each of the primary and secondary outcomes. Separate analyses were run for 4-week and 6-month outcomes comparing initial (pre-treatment) values with 4-week and 6-month values respectively.

Treatment response was investigated by determining whether changes in fear-avoidance beliefs, pain catastrophizing, and physical impairment were associated with changes in primary outcome measures. Residualized change scores were used for this part of the analysis because they control for the correlation between pre-treatment and post-treatment scores, and the subsequent potential for raw difference scores to be unreliable and susceptible to error. Residualized change scores were created by regressing post-treatment scores on pre-treatment scores, and then the resultant residuals were used in subsequent analyses. Then, residualized change scores were used in multiple regression to examine treatment change based on our a priori hypotheses that change in psychological factors would be associated with change in disability and pain intensity.

Results

Figure 1 describes the number of patients screened for eligibility, excluded from the study, and eventually enrolled in the trial. The majority of exclusions were patients with symptom duration of greater than 24 weeks (73.0%). Patients that were eligible, but refused to participate in the study ($n = 25$) were of similar sex distribution, age, and duration of symptoms when compared to those that enrolled in the study ($p > 0.05$). Randomization resulted in 36 patients assigned to TBC, 37 to GA, and 35 to GX. Intervention was not considered allocated for 2 patients in each group. The most common reason for intervention not being allocated was attending less than 2 physical therapy sessions ($n = 5$), while one patient required surgery for a shoulder blood clot discovered after consenting to the study ($n = 1$). These exclusions left 102 patients with the intervention appropriately allocated and included in subsequent analyses. There were no adverse events reported during this trial.

The baseline, post-randomization descriptive statistics are summarized in Table 3. The groups were balanced on most variables, but post-randomization differences existed for duration of symptoms and pain catastrophizing ($p < 0.05$). Also, post-randomization differences existed if only the 2 most common TBC subgroups (mobilization and immobilization) were compared ($p < 0.05$). Neither duration nor pain catastrophizing was significantly correlated with the primary outcome measures. Furthermore, the mobilization and immobilization subgroups did not differ on changes in primary outcome measures. Thus these variables did not meet the other assumption of covariate analysis and were not included in subsequent models despite post-randomization differences. The treatment parameters, follow-up rates, and follow-up times were similar across the intervention groups at 4-weeks and 6-months (Table 4). Furthermore, patients reported similar degrees of additional health care utilization and patient satisfaction at the 6-month assessment (Table 4).

Primary Outcomes

The 4-week and 6-month outcomes for disability and pain intensity are reported in Table 5. For disability only main effects for time were detected at 4-weeks [$F(1,86) = 59.5$, $p < .01$, $\eta^2 = .41$] and 6-months [$F(1,66) = 71.5$, $p < .01$, $\eta^2 = .52$] indicating all treatment groups showed a significant decrease in ODQ scores. The results for pain intensity ratings mirrored the findings for disability. Only main effects for time were detected at 4-weeks [$F(1,86) = 84.1$, $p < .01$, $\eta^2 = .49$] and 6-months [$F(1,66) = 100.2$, $p < .01$, $\eta^2 = .60$] indicating all treatment groups reported a significant decrease in NRS ratings. Specific to our primary hypothesis, GX was not more effective than GA or TBC for reducing disability and pain intensity for patients

with elevated fear-avoidance beliefs (Figure 2). Effect sizes (Cohen's *d*) of the planned comparisons (involving only patients with elevated fear-avoidance beliefs) were calculated to allow additional understanding of the magnitude of these effects. At 4-weeks the effect sizes for ODQ scores were GX vs. GA (-.40), GX vs. TBC (-.02), and GA vs. TBC (.39) and for pain intensity were GX vs. GA (.11), GX vs. TBC (-.05), and GA vs. TBC (-.16). At 6-months the effect sizes for ODQ scores were GX vs. GA (-.38), GX vs. TBC (-.37), and GA vs. TBC (.01) and for pain intensity were GX vs. GA (-.32), GX vs. TBC (-.26), and GA vs. TBC (.01). These post-hoc effect sizes suggest that for the primary comparisons of interest (GX vs. GA and GX vs. TBC) total sample sizes needed to detect these magnitudes of differences would range from 114 to over 700.

Secondary Outcomes

The 4-week and 6-month outcomes for fear-avoidance beliefs, pain catastrophizing, and physical impairment are reported in Table 5. For fear-avoidance beliefs, only a main effect for time was observed at 4-weeks [$F(1,89) = 67.9, p < .01, \eta^2 = .43$] indicating all groups had a significant reduction in FABQ-PA scores. The main effect of time remained significant at 6-months [$F(1,69) = 55.2, p < .01, \eta^2 = .44$] indicating all groups had lower FABQ-PA scores. However, a treatment group \times time interaction was also significant at 6-months [$F(2,69) = 3.9, p = .02, \eta^2 = .10$]. Follow-up simple contrasts indicated that both TBC ($t = 5.3, p < .01, \text{Cohen's } d = 1.03$) and GX [$t = 5.5, p < .01, \text{Cohen's } d = 1.12$] showed significant decreases in pain-related fear while GA [$t = 2.1, p = .05, \text{Cohen's } d = .40$] had a smaller treatment response (Figure 2).

For pain catastrophizing, main effects for time were observed at 4-weeks [$F(1,86) = 13.1, p < .01, \eta^2 = .13$] and 6-months [$F(1,69) = 6.8, p = .01, \eta^2 = .09$] suggesting all groups showed similar decrease in PCS scores. At 4-weeks [$F(2,89) = 4.4, p = .02, \eta^2 = .09$] and 6-months [$F(2,69) = 3.9, p = .02, \eta^2 = .10$] a main effect of treatment group was significant suggesting that the GX group had significantly higher overall PCS scores. This was consistent with the baseline post-randomization analysis already reported and suggests these differences persisted throughout the trial. No treatment group \times time interactions reached statistical significance in the model for pain catastrophizing. For physical impairment, only a main effect for time was observed at 4-weeks [$F(1,74) = 94, p < .01, \eta^2 = .56$] was observed, suggesting all treatment groups showed a similar decrease in PIS scores following treatment.

Treatment Response

4-week ODQ residualized change scores indicated that the combination of FABQ-PA, PIS, PCS, and NRS residualized change scores accounted for 47% of the variance ($p < .001$). At 4-weeks only pain intensity change ($\beta = .58, p < .01$) was uniquely associated with change in disability, not changes in fear-avoidance beliefs ($\beta = .07$), pain catastrophizing ($\beta = .16$), and physical impairment ($\beta = .09$). 4-week NRS residualized change scores using FABQ-PA, PIS, and PCS residualized change scores accounted for 23% of the variance ($p < .01$). At 4-weeks changes in both fear-avoidance beliefs ($\beta = .26, p = .01$) and physical impairment ($\beta = .35, p < .01$) were uniquely associated with change in pain intensity, not change in pain catastrophizing ($\beta = .11$).

At 6-months the combination of residualized change in NRS, PCS, and FABQ-PA scores accounted for 65% of the variance in ODQ residualized change ($p < .01$). The standardized beta weights indicated that pain intensity ($\beta = .70, p < .01$) was uniquely associated with 6-month disability change, not fear-avoidance beliefs ($\beta = .10$) and pain catastrophizing ($\beta = .10$). Regression results indicated that FABQ-PA and PCS residualized change scores accounted for NRS residualized change scores at 6-months ($R^2 = .28, p < .01$). Changes in both

fear-avoidance beliefs ($\beta = .25$, $p = .02$) and pain catastrophizing ($\beta = .38$, $p < 0.01$) were uniquely associated with change in pain intensity scores.

Categorical Outcomes

Previous literature has reported cutoffs scores to define clinically meaningful changes for the ODQ (greater than 10 points)[34], NRS (greater than 2 points)[6], and FABQ-PA (greater than 4 points)[4]. Also, it has been suggested that estimates of treatment effect include categorical outcomes for clinical trials involving pain research[10]. Therefore, proportion of 6-month success vs. failure rates were calculated and compared across treatment groups. For ODQ, the 6-month success rates were 56%, 41%, 43% for TBC, GA, and GX respectively. The success rate for ODQ change was not significantly different between groups ($p = .70$). For pain intensity, the 6-month success rates were 42%, 43%, and 46% for TBC, GA, and GX, respectively. The success rate for pain intensity change did not differ between treatment groups ($p = .76$). For FABQ-PA, the 6-month success rates were 42%, 22%, and 43% for TBC, GA, and GX, respectively. The success rate for FABQ-PA change did not differ between groups ($p = .17$).

Intention to Treat Analysis

Baseline age, sex, disability, pain intensity, fear-avoidance beliefs, pain catastrophizing, and physical impairment were compared between those who completed the 4-week and 6-month assessments and those who did not. None of the variables showed statistically significant ($p > .05$) differences for these comparisons. Then, all primary outcome variables were subjected to an ITT analysis where pre-treatment values were substituted for follow-up values when missing. Results of these analyses showed continued main effects for time on both ODQ and pain intensity with none of the hypothesized interaction effects. These ITT analyses were virtually identical to the completers only analyses already reported, so the ITT analyses are not repeated here. Collectively these analyses indicate that dropouts should not adversely influence the interpretation of these data.

Discussion

The purpose of this clinical trial was to compare the effectiveness of 3 different forms of physical therapy (TBC, GA, and GX) for patients with acute and sub-acute LBP. Our primary hypothesis was that patients with elevated pain related fear would benefit most from GX. This hypothesis was not supported as only main effects for time were observed for primary outcomes. Effect sizes indicated poorest disability and pain outcomes for GX; however these findings were not statistically significant. In our previous trial disability scores were higher and GA was more effective than TBC for those with elevated fear-avoidance beliefs. However, these results were not replicated in this trial. Our hypothesis was supported for one secondary outcome (fear-avoidance beliefs) as GX had better FABQ-PA outcomes at 6-months, in comparison to GA. TBC had similar fear-avoidance beliefs outcomes as GX which was an unexpected finding. The current findings suggest that supplementing TBC with GA or GX was not effective for improving important outcomes related to the development of chronic LBP.

Recently reported clinical trials provide comparison data for patients with chronic LBP. Smeets et al [38] reported that cognitive-behavioral treatment, active physical treatment, and a combination of the 2 treatments resulted in similar clinical outcomes for single or combined treatments. Linton et al[30] reported that GX was more effective than a wait list control group receiving usual care for improvements in function, but not for pain intensity or pain-related fear. Leeuw et al [27] reported no statistically significant differences between GA and GX for functional disability, main complaints, daily activity levels, or pain intensity. Consistent with the primary hypothesis of this trial, Leeuw et al[27] reported that the effects of GX and GA

were equivalent for patients with elevated pain related fear. Although this literature is relatively small, the available data from previous trials[27,30] and our own trial suggest that GX has the potential to be effective in comparison to usual care, but not in comparison to GA.

TBC differs from typical administration of physical therapy in clinical trials of LBP and this is an important factor to consider when interpreting these results. Our previous clinical trial investigating GA included TBC using different guidelines based on clinical expertise[18]. In the current trial empirically generated clinical prediction rules guided treatment decisions[5, 12,20]. These clinical prediction rules utilized different factors for treatment selection in comparison to previous TBC guidelines. Therefore, treatment was linked to factors previously associated with a higher probability of a positive outcome. Three recent clinical trials have reported that treatment guided by TBC and the empirically generated clinical prediction rules was more effective for LBP than guideline based treatment[16], unmatched treatment[3], or stabilization exercises[5]. There is mounting evidence that TBC is an effective form of physical therapy for LBP. The robust treatment effects associated with empirically generated clinical prediction rules are a potential explanation for why neither GA nor GX supplements resulted in improved clinical outcomes.

Another interpretation of these data is ineffective forms of GA and GX were utilized. We do not suspect this to be a concern for GA as we have reported previous evidence for its effectiveness for those with elevated pain related fear[18], and the same guidelines were used in this trial. The implementation of GX was by a self-report questionnaire to identify fearful activities. This procedure differs from what recently has been described as state of the art in the literature, the Photograph Series of Daily Activities (PHODA)[26]. We utilized a pragmatic approach for identifying fearful activities because we needed a process that would minimally disrupt the clinical environment and the shortened version of the PHODA was unavailable at the start of this trial. The empirical support for successful implementation of GX in this trial comes from the lower 6-month fear-avoidance beliefs. While these results are encouraging, additional data are necessary to determine if self-report questionnaires are an acceptable way to implement GX.

There are obvious theoretical differences in GX and GA for determining exercise and activity parameters. Those differences were accounted for when creating the dosing paradigms utilized in this study. However, implementation of GX and GA may result in treatment overlap and this is another potential explanation for the lack of differences in the behavioral supplements. First, both approaches utilized an education approach of reducing the fear and threat associated with LBP. Second, dosing exercise for GA is not absent of exposure to exercise and activities of which the patient might also be fearful. Activities that were commonly rated as fearful for subjects participating in this study (i.e. lifting, carrying, and reaching to the floor) would typically be included in GA paradigms. Our current data suggest that it may not be necessary to implement a formal exposure paradigm for patients with sub-acute or acute LBP if the GA component includes a variety of activities that subjects are likely to fear.

The use of acute and sub-acute patients with LBP could be another reason the 3 treatment groups had similar outcomes. A clinical trial in primary practice found no benefit from a minimal intervention targeted at psychosocial prognostic factors for patients with sub-acute LBP[22]. We targeted patients with acute and sub-acute LBP because one of our goals was to determine if a more intensive intervention like GX was effective for secondary prevention of LBP. Potentially effective secondary prevention models have been reported by Linton et al [29,31]. Therefore, it seemed reasonable to hypothesize that GX had the potential to reduce the likelihood of transitioning to chronic LBP. GX was not more effective in the current study, but the Linton et al studies [29,31] included patients with neck and LBP, were performed in occupational settings, utilized different interventions, and included sick leave and health care

utilization as primary outcome measures. Future studies are necessary to clarify these methodological differences, and to determine the optimal treatment approach to prevent patients with acute and sub-acute symptoms from transitioning to chronic pain syndromes.

As expected, disability reduction was associated with reductions in pain intensity and physical impairment at 4-weeks, and reduction in pain intensity alone at 6-months. Reductions in fear-avoidance beliefs and pain catastrophizing did not significantly contribute to the reduction of disability at 4-weeks or 6-months. These findings differ from previous studies demonstrating that reduction in psychological factors were associated with reduction in disability, even after controlling for pain intensity[19,39,52]. Pain intensity reduction was associated with reductions in fear-avoidance beliefs and pain catastrophizing. These data are consistent with the previously mentioned studies of chronic[39,52] and sub-acute LBP[19]. The immediate clinical implication for patients with acute and sub-acute LBP seems to be that reduction of these psychological variables remains an appropriate treatment target because of their association with pain intensity. However, reduction of pain intensity and physical impairment remains the primary goal for patients with acute and sub-acute LBP. The novel information gained from this randomized trial is that it may not be necessary to apply GX or GA to reduce psychological variables if TBC is used.

The primary limitation of this study is the lack of a true or wait list control group, neither of which was an option in this clinical setting. Therefore, we can not make conclusions about absolute treatment effectiveness for TBC, GA, or GX. We powered this study to detect a treatment effect of 5-7% and reached our targeted sample size. Post-hoc calculations of effect sizes indicated that future studies will need much larger sample sizes to detect statistically significant group differences. Even with larger samples, however, the clinical importance of similar magnitude findings is debatable[10]. Another limitation is that we only used pain related fear to determine our groups for being at risk of developing chronic LBP to keep the methodology consistent with our previous trial[18]. Other research has suggested that mood, stress, or pain catastrophizing comprise musculoskeletal pain subgroups with increased risk of developing chronic symptoms[2,32]. Other limitations include the existence of post-randomization differences for duration of symptoms, pain catastrophizing, and 2 of the TBC groups. Our covariate analysis suggests minimal influence on outcomes, but these differences could still have influenced study results. Last, we assessed if LBP was work related, but did not assess whether litigation was involved, which could be considered another limitation.

Future study in patients with LBP should focus on improving the identification of patients at risk for developing chronic LBP and refining specific psychological targets that optimally reduce pain intensity in acute and sub-acute phase[21]. Future studies should also define optimal clinical levels of reduction for these psychological treatment targets. Future study could consider the use of phobia criteria to determine a subgroup most likely to need GX. Using standard phobia criteria coupled with specific phobic stimulus exposure is likely to yield the highest treatment effect in future studies. Last, future study should also determine the relative effectiveness of GX and GA based on stage of LBP and include cost data. Currently there is not much data to support that these treatments differ largely in their effectiveness, so timing and cost of the intervention may be important considerations for future clinical application.

Acknowledgements

The authors of this manuscript have no financial or other relationship with any individuals or organizations that constitute a conflict of interest.

SZG (PI), MER, and CV were supported by NIH-NIAMS grant AR051128 while preparing this manuscript.

SHANDS Rehab at the University of Florida allowed for utilization of facilities.

Tim Day, Kevin McDonald, and Pat O'Connor provided administrative support.

Reference List

1. Scientific approach to the assessment and management of activity-related spinal disorders. A monograph for clinicians. Report of the Quebec Task Force on Spinal Disorders. *Spine* 1987;12:S1–59. [PubMed: 2961086]
2. Boersma K, Linton SJ. Screening to identify patients at risk: profiles of psychological risk factors for early intervention. *Clin J Pain* 2005;21:38–43. [PubMed: 15599130]
3. Brennan GP, Fritz JM, Hunter SJ, Thackerary A, Delitto A, Erhard RE. Identifying sub-groups of patients with “non-specific” low back pain: results of a randomized clinical trial. *Spine*. 2005in press
4. Burton AK, Waddell G, Tillotson KM, Summerton N. Information and advice to patients with back pain can have a positive effect. A randomized controlled trial of a novel educational booklet in primary care. *Spine* 1999;24:2484–2491. [PubMed: 10626311]
5. Childs JD, Fritz JM, Flynn TW, Irrgang JJ, Johnson KK, Majkowski GR, et al. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. *Ann Intern Med* 2004;141:920–928. [PubMed: 15611489]
6. Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. *Spine* 2005;30:1331–1334. [PubMed: 15928561]
7. Daltroy LH, Cats-Baril WL, Katz JN, Fossel AH, Liang MH. The North American spine society lumbar spine outcome assessment Instrument: reliability and validity tests. *Spine* 1996;21:741–749. [PubMed: 8882698]
8. Delitto A, Erhard RE, Bowling RW. A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment [see comments]. *Physical Therapy* 1995;75:470–485. [PubMed: 7770494]
9. Deyo RA, Battie M, Beurskens AJ, Bombardier C, Croft P, Koes B, et al. Outcome measures for low back pain research. A proposal for standardized use. *Spine* 1998;23:2003–2013. [PubMed: 9779535]
10. Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J Pain* 2008;9:105–121. [PubMed: 18055266]
11. Fairhurst K, Dowrick C. Problems with recruitment in a randomized controlled trial of counselling in general practice: causes and implications. *J Health Serv Res Policy* 1996;1:77–80. [PubMed: 10180853]
12. Flynn T, Fritz J, Whitman J. A clinical prediction rule for classifying patients with low back pain who demonstrate short term improvement with spinal manipulation. *Spine*. 2002in press
13. Fordyce WE, Fowler RS, Lehmann JF, Delateur BJ, Sand PL, Trieschmann RB. Operant conditioning in the treatment of chronic pain. *Arch Phys Med Rehabil* 1973;54:399–408. [PubMed: 4729785]
14. Frank JW, Brooker AS, DeMaio SE, Kerr MS, Maetzel A, Shannon HS, et al. Disability resulting from occupational low back pain. Part II: What do we know about secondary prevention? A review of the scientific evidence on prevention after disability begins. *Spine* 1996;21:2918–2929. [PubMed: 9112717]
15. Fritz JM, Brennan GP, Clifford SN, Hunter SJ, Thackeray A. An examination of the reliability of a classification algorithm for subgrouping patients with low back pain. *Spine* 2006;31:77–82. [PubMed: 16395181]
16. Fritz JM, Delitto A, Erhard RE. Comparison of classification-based physical therapy with therapy based on clinical practice guidelines for patients with acute low back pain: a randomized clinical trial. *Spine* 2003;28:1363–1371. [PubMed: 12838091]
17. Fritz JM, George S. The use of a classification approach to identify subgroups of patients with acute low back pain. Interrater reliability and short-term treatment outcomes. *Spine* 2000;25:106–114. [PubMed: 10647168]
18. George SZ, Fritz JM, Bialosky JE, Donald DA. The effect of a fear-avoidance-based physical therapy intervention for patients with acute low back pain: results of a randomized clinical trial. *Spine* 2003;28:2551–2560. [PubMed: 14652471]

19. George SZ, Fritz JM, McNeil DW. Fear-avoidance beliefs as measured by the fear-avoidance beliefs questionnaire: change in fear-avoidance beliefs questionnaire is predictive of change in self-report of disability and pain intensity for patients with acute low back pain. *Clin J Pain* 2006;22:197–203. [PubMed: 16428956]
20. Hicks GE, Fritz JM, Delitto A, McGill SM. Preliminary development of a clinical prediction rule for determining which patients with low back pain will respond to a stabilization exercise program. *Arch Phys Med Rehabil* 2005;86:1753–1762. [PubMed: 16181938]
21. Jellema P, van der Windt DA, van der Horst HE, Blankenstein AH, Bouter LM, Stalman WA. Why is a treatment aimed at psychosocial factors not effective in patients with (sub)acute low back pain? *Pain* 2005;118:350–359. [PubMed: 16289797]
22. Jellema P, van der Windt DA, van der Horst HE, Twisk JW, Stalman WA, Bouter LM. Should treatment of (sub)acute low back pain be aimed at psychosocial prognostic factors? Cluster randomised clinical trial in general practice. *BMJ* 2005;331:84. [PubMed: 15967762]
23. Jensen MP, Turner JA, Romano JM, Fisher LD. Comparative reliability and validity of chronic pain intensity measures. *Pain* 1999;83:157–162. [PubMed: 10534586]
24. Jensen MP, Turner LR, Turner JA, Romano JM. The use of multiple-item scales for pain intensity measurement in chronic pain patients. *Pain* 1996;67:35–40. [PubMed: 8895229]
25. Leeuw M, Goossens ME, Linton SJ, Crombez G, Boersma K, Vlaeyen JW. The fear-avoidance model of musculoskeletal pain: current state of scientific evidence. *J Behav Med* 2007;30:77–94. [PubMed: 17180640]
26. Leeuw M, Goossens ME, van Breukelen GJ, Boersma K, Vlaeyen JW. Measuring perceived harmfulness of physical activities in patients with chronic low back pain: the Photograph Series of Daily Activities--short electronic version. *J Pain* 2007;8:840–849. [PubMed: 17632038]
27. Leeuw M, Goossens ME, van Breukelen GJ, de J, Heuts PH, Smeets RJ, et al. Exposure in vivo versus operant graded activity in chronic low back pain patients: Results of a randomized controlled trial. *Pain*. 2008
28. Lindstrom I, Ohlund C, Eek C, Wallin L, Peterson LE, Fordyce WE, et al. The effect of graded activity on patients with subacute low back pain: a randomized prospective clinical study with an operant-conditioning behavioral approach. *Phys Ther* 1992;72:279–290. [PubMed: 1533941]
29. Linton SJ, Andersson T. Can chronic disability be prevented? A randomized trial of a cognitive-behavior intervention and two forms of information for patients with spinal pain. *Spine* 2000;25:2825–2831. [PubMed: 11064530]
30. Linton SJ, Boersma K, Jansson M, Overmeer T, Lindblom K, Vlaeyen JW. A randomized controlled trial of exposure in vivo for patients with spinal pain reporting fear of work-related activities. *Eur J Pain*. 2007
31. Linton SJ, Boersma K, Jansson M, Svard L, Botvalde M. The effects of cognitive-behavioral and physical therapy preventive interventions on pain-related sick leave: a randomized controlled trial. *Clin J Pain* 2005;21:109–119. [PubMed: 15722803]
32. Linton SJ, Hallden K. Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain. *Clin J Pain* 1998;14:209–215. [PubMed: 9758070]
33. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *J Am Podiatr Med Assoc* 2001;91:437–442. [PubMed: 11574648]
34. Ostelo RW, de Vet HC. Clinically important outcomes in low back pain. *Best Pract Res Clin Rheumatol* 2005;19:593–607. [PubMed: 15949778]
35. Pincus T, Burton AK, Vogel S, Field AP. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine* 2002;27:E109–E120. [PubMed: 11880847]
36. Pincus T, Vlaeyen JW, Kendall NA, Von Korff MR, Kalauokalani DA, Reis S. Cognitive-behavioral therapy and psychosocial factors in low back pain: directions for the future. *Spine* 2002;27:E133–E138. [PubMed: 11880850]
37. Roland, M.; Waddell, G.; Klaber-Moffett, J. *The Back Book*. Norwich, United Kingdom: The Stationery Office; 1996. Ref Type: Pamphlet

38. Smeets RJ, Vlaeyen JW, Hidding A, Kester AD, van der Heijden GJ, van Geel AC, et al. Active rehabilitation for chronic low back pain: Cognitive-behavioral, physical, or both? First direct post-treatment results from a randomized controlled trial [ISRCTN22714229]. *BMC Musculoskelet Disord* 2006;7:5. [PubMed: 16426449]
39. Smeets RJ, Vlaeyen JW, Kester AD, Knottnerus JA. Reduction of pain catastrophizing mediates the outcome of both physical and cognitive-behavioral treatment in chronic low back pain. *J Pain* 2006;7:261–271. [PubMed: 16618470]
40. Smeets RJ, Wittink H. The deconditioning paradigm for chronic low back pain unmasked? *Pain* 2007;130:201–202. [PubMed: 17532142]
41. Sullivan MJL, Bishop SR, Pivik J. The Pain Catastrophizing Scale: development and validation. *Psychological Assessment* 1995;7:524–532.
42. Torgerson DJ, Klaber-Moffett J, Russell IT. Patient preferences in randomised trials: threat or opportunity? *J Health Serv Res Policy* 1996;1:194–197. [PubMed: 10180870]
43. Turk DC. The potential of treatment matching for subgroups of patients with chronic pain: lumping versus splitting. *Clin J Pain* 2005;21:44–55. [PubMed: 15599131]
44. Vlaeyen JW, de Jong J, Geilen M, Heuts PH, Van Breukelen G. Graded exposure in vivo in the treatment of pain-related fear: a replicated single-case experimental design in four patients with chronic low back pain. *Behav Res Ther* 2001;39:151–166. [PubMed: 11153970]
45. Vlaeyen JW, de Jong J, Geilen M, Heuts PH, Van Breukelen G. The treatment of fear of movement/ (re)injury in chronic low back pain: further evidence on the effectiveness of exposure in vivo. *Clin J Pain* 2002;18:251–261. [PubMed: 12131067]
46. Vlaeyen JW, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. *Pain* 2000;85:317–332. [PubMed: 10781906]
47. Von Korff M, Balderson BH, Saunders K, Miglioretti DL, Lin EH, Berry S, et al. A trial of an activating intervention for chronic back pain in primary care and physical therapy settings. *Pain* 2005;113:323–330. [PubMed: 15661440]
48. Waddell G. 1987 Volvo award in clinical sciences. A new clinical model for the treatment of low-back pain. *Spine* 1987;12:632–644. [PubMed: 2961080]
49. Waddell G. Low back disability. A syndrome of Western civilization. *Neurosurg Clin N Am* 1991;2:719–738. [PubMed: 1840384]
50. Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain* 1993;52:157–168. [PubMed: 8455963]
51. Waddell G, Somerville D, Henderson I, Newton M. Objective clinical evaluation of physical impairment in chronic low back pain. *Spine* 1992;17:617–628. [PubMed: 1308095]
52. Woby SR, Watson PJ, Roach NK, Urmston M. Are changes in fear-avoidance beliefs, catastrophizing, and appraisals of control, predictive of changes in chronic low back pain and disability? *Eur J Pain* 2004;8:201–210. [PubMed: 15109970]

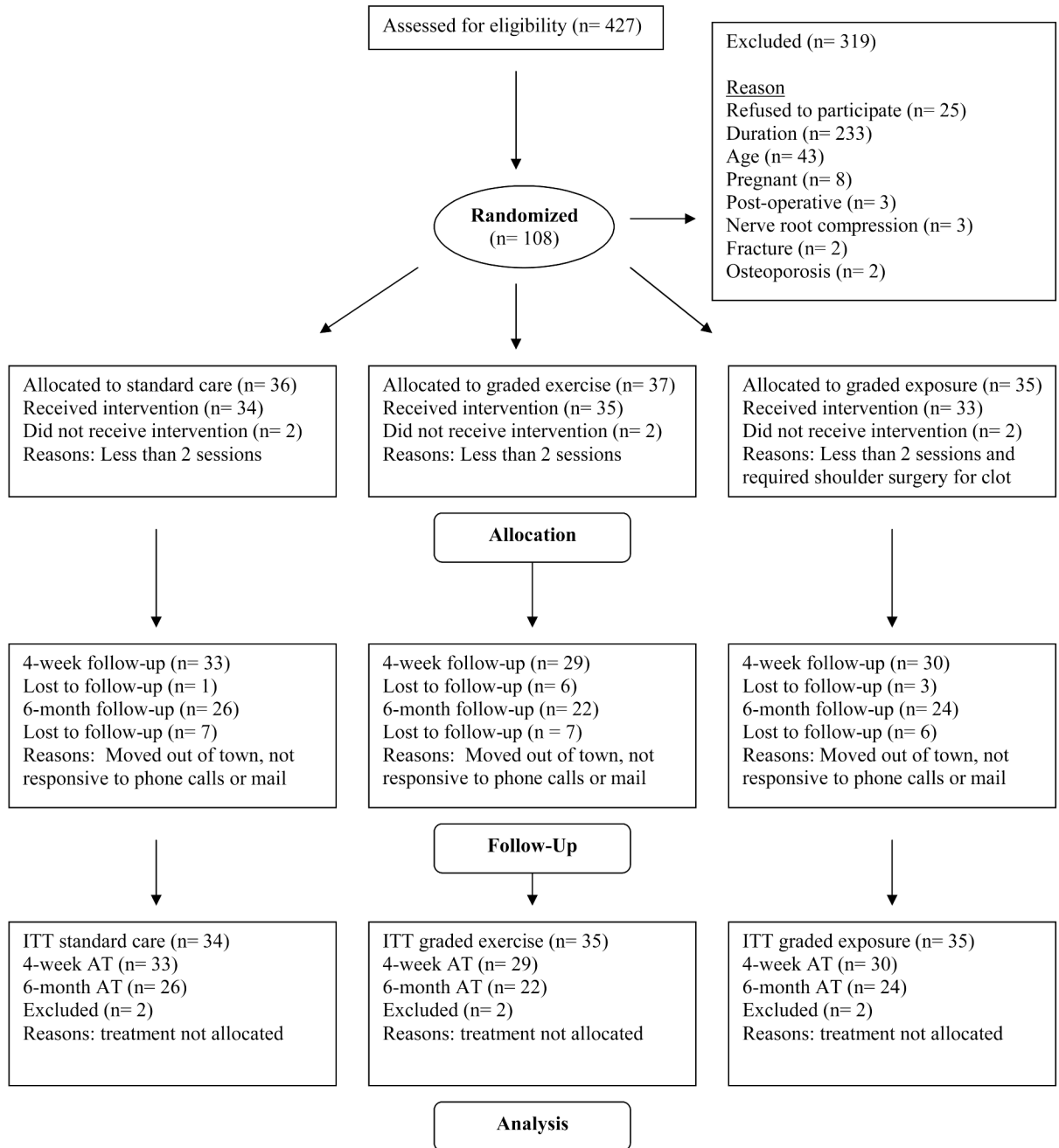


Figure 1. Summary of Recruitment, Enrollment, Allocation, Follow-up, and Analysis for Behavioral Intervention Trial

ITT – Intention to treat analysis, AT – As treated (completers only) analysis

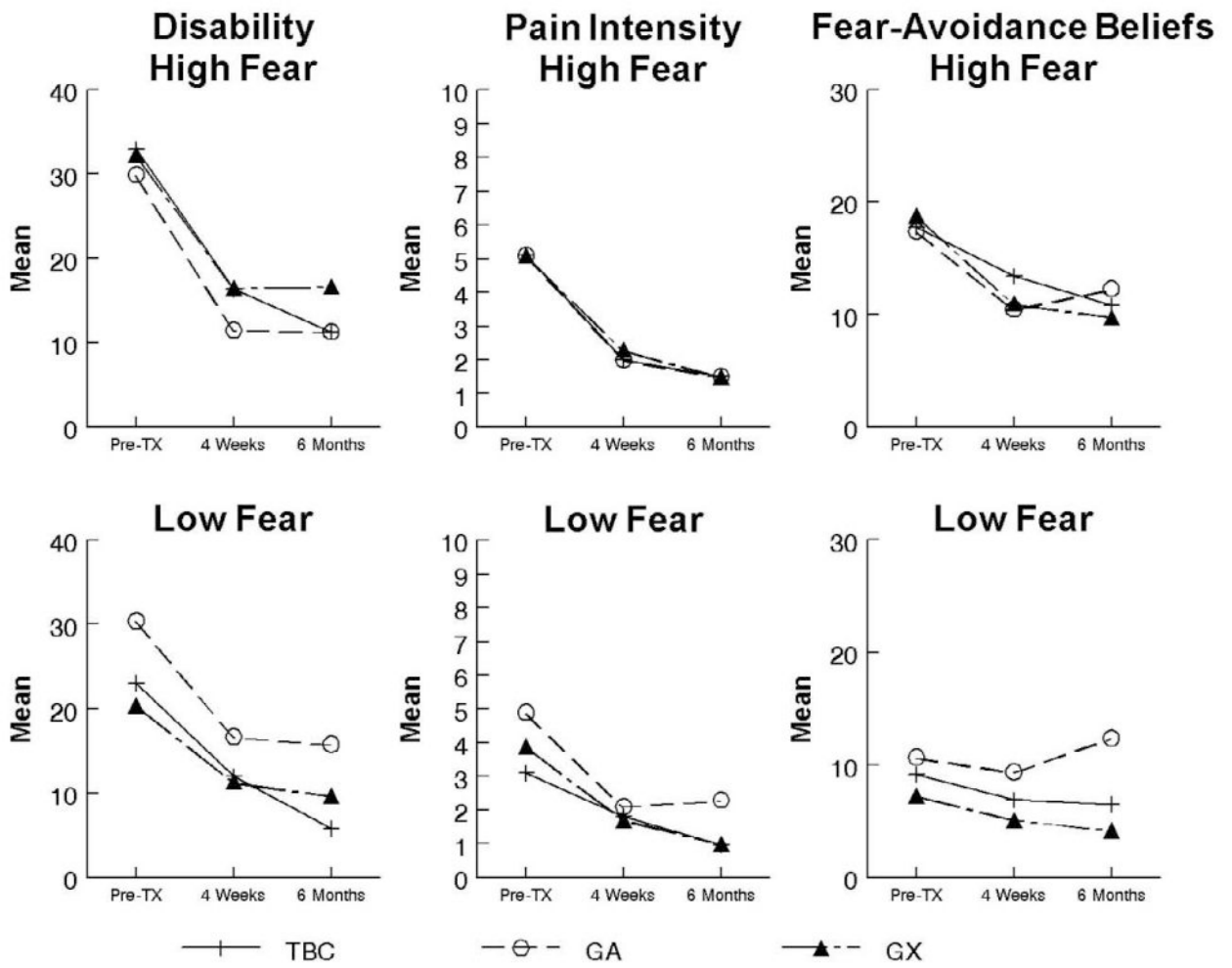


Figure 2. Comparison of Disability, Pain Intensity, and Fear-Avoidance Outcomes by Initial Pain Related Fear Level

Table 1
Key Differences in Interventions used in Clinical Trial

	Treatment Based Classification	Graded Activity	Graded Exposure
Treatment matched by empirically generated clinical prediction rule?	Yes	Yes	Yes
Treatment progressed by pain abatement?	Yes	No	No
Treatment progressed by quota based system for therapeutic exercise and general activity tolerance?	No	Yes	No
Treatment progressed by specifically performing fearful activities?	No	No	Yes

Table 2
Summary of Fear of Daily Activities Ratings for All Subjects Participating in Clinical Trial

Activity	Mean (SD)	25 th Percentile	50 th Percentile	75 th Percentile
Sitting for longer than 1 hour	39.1 (34.8)	2	40	71
Standing for longer than 30 minutes	37.0 (35.6)	0	28	70
Walking for longer than 30 minutes	35.4 (35.1)	0	35	66
Lifting less than 20 pounds	26.3 (30.1)	0	13	50
Lifting more than 20 pounds	50.6 (35.6)	20	50	85
Carrying less than 20 pounds	25.3 (28.8)	0	15	43
Carrying more than 20 pounds	52.0 (35.3)	20	60	88
Twisting	38.4 (32.3)	10	40	60
Reaching to floor	41.3 (34.1)	9	40	66
Performing back exercises	35.6 (32.7)	5	28	53

Numerical scale ratings ranging from 0 (no fear) to 100 (worst fear imaginable) for each activity. Activities with highest individual ratings were implemented in graded exposure protocol.

Table 3
Baseline Characteristics for Subjects Enrolled in Behavioral Interventions Trial

Variable	Treatment Based Classification (n = 34)	Graded Activity (n = 35)	Graded Exposure (n = 33)	p-value
Age [*]	34.9 (14.4)	37.6 (15.8)	40.1 (14.5)	.346
Sex [^] (# of female subjects)	23 (68%)	24 (69%)	21 (64%)	.901
Currently employed (# employed)	25 (74%)	26 (74%)	23 (70%)	.845
Work related LBP (# with work related LBP)	6 (18%)	7 (20%)	8 (24%)	.971
Marital status [^] (# of married subjects)	13 (38%)	17 (49%)	15 (45%)	.676
Education [^] (# of college graduates)	21 (62%)	16 (46%)	14 (42%)	.268
Prior history of LBP [^] (# with prior history)	17 (50%)	24 (69%)	22 (67%)	.288
Leg pain with LBP [^] (# with leg pain)	13 (38%)	17 (49%)	14 (42%)	.787
Onset of LBP [^] (# with sudden onset)	23 (68%)	24 (69%)	21 (64%)	.804
Duration of LBP [^] (# of weeks present episode)	6.7 (4.8)	5.8 (5.5)	9.8 (7.0)	.015
Treatment preference				.849
Physical therapy only	5 (15%)	5 (14%)	7 (21%)	
Graded exercise	9 (26%)	15 (43%)	12 (36%)	
Graded exposure	3 (9%)	2 (6%)	2 (6%)	
No preference	15 (44%)	12 (34%)	12 (36%)	
Treatment classification [^]				.123
Specific exercise	3 (9%)	2 (6%)	3 (9%)	
Mobilization [#]	18 (53%)	20 (57%)	8 (24%)	
Immobilization [#]	10 (29%)	12 (34%)	20 (61%)	
Traction	1 (3%)	1 (3%)	0 (0%)	
Unable to classify	2 (6%)	0 (0%)	2 (6%)	
Oswestry disability score [*]	29.2 (15.7)	31.1 (15.8)	30.7 (15.6)	.346
Average pain intensity [*]	4.3 (2.0)	5.2 (1.8)	4.7 (2.1)	.215
Fear-avoidance beliefs about physical activity [*]	14.9 (4.8)	15.0 (4.4)	16.1 (6.0)	.558
Fear-avoidance beliefs about work [*]	12.5 (10.1)	15.8 (12.5)	13.9 (11.5)	.487
Pain catastrophizing [*]	12.6 (9.0)	16.4 (11.3)	20.7 (11.3)	.025
Physical impairment [*]	2.9 (1.7)	3.6 (2.1)	3.1 (1.6)	.338

All values reported as mean (standard deviation) or number (percentage within treatment assignment)

* - Significance tested with ANOVA

[^] - Significance tested with chi-square test

[#] - When only mobilization and immobilization are compared there is significant difference ($p < 0.05$)

Table 4
Treatment Summary for Behavioral Interventions Trial

Four Weeks	Treatment Based Classification	Graded Activity	Graded Exposure	p-value
Days to follow-up [*]	29.6 (6.2)	30.6 (6.6)	30.6 (10.2)	.853
Follow-up completed [^]	33 (97%)	29 (83%)	30 (91%)	.138
Number of appointments [*]				.116
Mean	6.2 (2.1)	5.4 (1.9)	6.2 (1.8)	
Median	6	5	6	
Range	2 – 12	2 – 9	2 – 10	
Six Months	Treatment Based Classification	Graded Activity	Graded Exposure	p-value
Days to follow-up [*]	181.5 (7.3)	183.7 (11.1)	184.7 (12.2)	.655
Follow-up completed [^]	26 (77%)	22 (63%)	24 (73%)	.439
Had other treatment since last research assessment	12 (55%)	10 (50%)	14 (61%)	.771
Satisfied with symptoms if present rest of life	13 (50%)	8 (36%)	12 (50%)	.564
Would select same treatment again	24 (92%)	20 (91%)	21 (88%)	.842
Overall rating of positive treatment effect	23 (89%)	19 (86%)	21 (88%)	.976
Positive rating of information	24 (92%)	21 (96%)	22 (92%)	.865

Unless otherwise indicated all values reported as mean (standard deviation) for continuous data or number (percentage within treatment assignment) for categorical data

* - Significance tested with ANOVA

[^] - Significance tested with chi-square test

Table 5
Descriptive Statistics for Primary and Secondary Outcomes

Outcome Measure	Group	Baseline High fear	Low fear	4-weeks* High Fear	Low Fear	6-months* High Fear	Low Fear
ODQ (0 – 100)	TBC	32.9 (16.1)	23.0 (15.5)	16.4 (14.9)	12.0 (11.5)	11.4 (11.5)	5.8 (7.1)
	GA	29.9 (18.4)	30.4 (13.3)	11.5 (11.8)	16.7 (11.9)	11.3 (14.2)	15.8 (11.1)
	GX	32.3 (16.3)	20.4 (13.1)	16.5 (12.1)	11.4 (11.6)	16.7 (17.6)	9.7 (8.2)
NRS (0 – 10)	TBC	5.1 (1.8)	3.1 (2.1)	2.0 (1.6)	1.8 (1.9)	1.6 (1.3)	1.0 (1.2)
	GA	5.1 (1.9)	4.9 (2.1)	2.3 (2.1)	2.1 (2.1)	1.5 (2.1)	2.3 (1.7)
	GX	5.1 (2.1)	3.9 (1.5)	2.1 (2.0)	1.7 (.9)	2.1 (2.3)	1.0 (1.0)
FABQ# (0 – 24)	TBC	17.8 (2.6)	9.2 (3.9)	13.4 (4.8)	6.9 (4.1)	10.8 (6.0)	6.5 (4.0)
	GA	17.4 (3.3)	10.7 (4.3)	10.5 (6.6)	9.4 (5.3)	12.3 (4.4)	12.4 (4.7)
	GX	18.8 (3.3)	7.3 (3.9)	11.0 (6.0)	5.1 (3.5)	9.8 (5.8)	4.2 (3.5)
PCS^ (0 – 52)	TBC	14.4 (9.6)	10.6 (9.8)	11.4 (8.5)	11.9 (9.7)	11.7 (11.2)	11.3 (7.7)
	GA	15.6 (11.8)	15.8 (11.3)	8.7 (7.1)	10.7 (8.5)	10.4 (9.5)	10.0 (5.2)
	GX	23.7 (9.9)	10.6 (10.4)	19.4 (8.2)	4.4 (5.4)	18.2 (14.9)	6.0 (10.9)
PIS (0 – 7)	TBC	3.6 (1.7)	2.1 (1.8)	1.3 (1.1)	.92 (1.3)		
	GA	3.5 (2.2)	3.5 (1.9)	1.1 (1.6)	1.7 (1.4)		
	GX	3.4 (1.9)	2.8 (1.8)	1.2 (1.3)	.86 (1.1)		

All values reported as mean (standard deviation) for completers of trial only

ODQ – Oswestry Disability Questionnaire, NRS – Numerical Rating Scale for pain intensity, FABQ – Fear Avoidance Beliefs Questionnaire (physical activity scale), PCS – Pain Catastrophizing Scale, and PIS – Physical Impairment Scale

* - Statistically significant main effects for time for all outcome measures ($p < 0.05$)

- Statistically significant treatment group \times time interaction at 6-months ($p < 0.05$)

^ - Statistically significant main effect for group at all times ($p < 0.05$)