Spine

SPINE Volume 42, Number 21, pp E1215–E1224

RANDOMIZED TRIAL

Who Benefits Most From Individualized Physiotherapy or Advice for Low Back Disorders? A Preplanned Effect Modifier Analysis of a Randomized Controlled Trial

Andrew J. Hahne, PhD, B.Physio(Hons),* Jon J. Ford, PhD, MPhysio BAppSc(Physio),* Matthew C. Richards, B.Physio,* Luke D. Surkitt, B.Physio,* Alexander Y.P. Chan, B.Physio (Hons),* Sarah L. Slater, BSc Physiotherapy (Hons), PhD,† and Nicholas F. Taylor, PhD, BAppSci(Physio) BSc*

Study Design. A preplanned effect modifier analysis of the Specific Treatment of Problems of the Spine randomized controlled trial.

Objective. To identify characteristics associated with larger or smaller treatment effects in people with low back disorders undergoing either individualized physical therapy or guideline-based advice.

Summary of Background Data. Identifying subgroups of people who attain a larger or smaller benefit from particular treatments has been identified as a high research priority for low back disorders.

Methods. The trial involved 300 participants with low back pain and/or referred leg pain (≥6 wk, ≤6 mo duration), who satisfied criteria to be classified into five subgroups (with 228 participants classified into three subgroups relating to disc-related disorders, and 64 classified into the zygapophyseal joint dysfunction subgroup). Participants were randomly allocated to receive either two sessions of guideline based advice (n = 144), or 10 sessions of individualized physical therapy targeting pathoanatomical, psychosocial, and neurophysiological factors (n = 156). Univariate and multivariate linear mixed models determined the interaction between treatment group and potential effect modifiers (defined a priori) for the primary outcomes of back pain, leg pain (0–10 Numeric Rating Scale) and activity limitation (Oswestry Disability Index) over a 52-week follow-up.

Results. Participants with higher levels of back pain, higher Örebro scores (indicative of higher risk of persistent pain) or longer duration of symptoms derived the largest benefits from individualized physical therapy relative to advice. Poorer coping also predicted larger benefits from individualized physical therapy in the univariate analysis.

Conclusion. These findings suggest that people with low back disorders could be preferentially targeted for individualized physical therapy rather than advice if they have higher back pain levels, longer duration of symptoms, or higher Örebro scores.

Key words: clinical response, low back pain, physical therapy, prediction.

Level of Evidence: 2

Spine 2017;42:E1215–E1224

From the *Low Back Research Team, School of Allied Health, La Trobe University, Melbourne, Australia; and †Physiotherapy Department, Northern Health, Melbourne, Australia.

Acknowledgment date: November 28, 2016. First revision date: January 15, 2017. Acceptance date: February 8, 2017.

The manuscript submitted does not contain information about medical device(s)/drug(s).

LifeCare Health was an industry partner that provided in-kind contribution (facilities, personnel, and resources) to allow treatment of participants free of charge. The sponsors had no role in the study planning, design, management, data analysis, writing of the manuscript, or the decision to submit the manuscript for publication.

Relevant financial activities outside the submitted work: employment, payment for development of educational presentations.

Address correspondence and reprint requests to Andrew J. Hahne, PhD, B.Physio (Hons), Low Back Research Team, School of Allied Health, La Trobe University, Kingsbury Dr, Melbourne, Victoria 3086, Australia; E-mail: a.hahne@latrobe.edu.au.

DOI: 10.1097/BRS.0000000000002148

Most randomized controlled trials (RCTs) relating to treatment of low back disorders demonstrate either no or modest effects.[1,2] Within RCTs there is typically considerable variation in how individuals respond to either the primary or comparison intervention.[3] Identifying subgroups of people who attain a larger or smaller treatment effect has been identified as a high research priority for low back disorders,[4] and has potential to assist health practitioners in selecting treatments most likely to be effective for a given person.[5]

There have been some attempts to identify subgroups who respond most favorably to particular treatments for low back disorders. Studies investigating treatment effect modifiers/moderators or clinical prediction rules require

large RCTs that are powered beyond the primary hypothesis of detecting overall treatment effects.[6,7] Perhaps due to this reason, systematic reviews have found that such studies are relatively scarce in the low back disorder literature, and those that have been undertaken are typically underpowered.[8–11]

The Specific Treatment of Problems of the Spine trial was an RCT that compared individualized physical therapy and advice to guideline-based advice alone for low back disorders.[12,13] The primary results showed that individualized physical therapy resulted in faster relief of back and leg pain, and sustained improvement in activity limitation (Oswestry) relative to guideline-based advice across a 12-month follow-up.[13] The aim of the present study was to investigate treatment effect modifiers in the Specific Treatment of Problems of the Spine trial to determine characteristics of participants who derived either a larger or smaller treatment effect from individualized physical therapy relative to advice. We hypothesized that participants with more severe or more complex low back disorders would derive the greatest benefit from individualized physical therapy relative to advice.

MATERIALS AND METHODS

The methods were based on guidelines for treatment effect modifier/moderator studies.[5,6,14,15] The data were obtained from a multicenter RCT comparing individualized physical therapy and advice (10 sessions) to guideline-based advice alone (two sessions) in 300 participants with low back disorders.[12,13] Informed consent was received from all participants before enrolment, and the study was approved by the La Trobe University Human Ethics Committee. The baseline characteristics of participants that were used for investigation of potential treatment effect modification were selected and declared a priori on the trial register (ACTRN12609000834257).

Participants

To be included in the trial, participants needed to have a current episode of low back pain (and/or referred leg pain) between 6 weeks and 6 months duration, be aged 18 to 65 years, speak English, and belong to one of five low back disorder subgroups being targeted in the trial (see Supplemental Digital Content, http://links.lww.com/BRS/B265). Exclusion criteria were a compensation claim, serious pathology (active cancer, cauda equine syndrome, foot drop making walking unsafe), pregnancy or childbirth within the last 6 months, history of lumbar spine surgery, spinal injections within the last 6 weeks, pain intensity less than 2/10 (leg pain for the disc herniation and associated radiculopathy subgroup, or back pain for the other subgroups) or minimal activity limitation.

Randomization and Concealment

Concealed random allocation of participants to either individualized physical therapy (n = 156) or advice (n = 144) was achieved *via* an offsite randomization service. Blinding of participants and therapists was not possible, but data entry was performed by a researcher who was blinded to treatment allocation.

Interventions

Treatment of both groups was delivered by 19 physical therapists across 16 private centers throughout metropolitan Melbourne, Australia.

Participants allocated to guideline-based advice received 2×30-minute sessions over a 10-week period based on the approach described by Indahl *et al*.[16] This included an explanation of the hypothesized pathoanatomical source of the participant's pain, reassurance regarding the likely favorable prognosis of their condition, advice to remain active, and instruction regarding lifting technique.[12]

Participants allocated to the individualized physical therapy group received 10×30-minute physical therapy sessions over a 10-week period. Physical therapy treatment was individualized firstly based on the five subgroups, four of which were pathoanatomical and one based on psychosocial factors (see Supplemental Digital Content, http://links.lww.com/BRS/B265). Further individualization was achieved within subgroups based on each participant's barriers to recovery. Available treatment components included pathoanatomical or neurophysiological information (matching the relevant subgroup), education (prognosis, treatment options), self-management strategies (posture, pacing, pain management, sleep management, relaxation strategies), inflammatory management, exercise rehabilitation (motor control training, goal-oriented functional exercises), manual therapy (zygapophyseal joint dysfunction subgroup only), and cognitive-behavioral strategies. Full details of the treatment protocols have been published previously.[12,17–20]

Baseline Predictors of Treatment Response

We prespecified ten potential treatment effect modifiers (Table 1). These features were assessed in all participants before randomization, eliminating the opportunity for assessor bias.[5,15]

Outcomes

Outcomes were obtained *via* postal questionnaire at baseline, and at 5, 10, 26, and 52 weeks after randomization. Activity limitation was assessed using the Oswestry Disability Index,[29] which has been shown to be valid and reliable for use in low back disorder populations.[30,31] Back pain and leg pain intensity were measured separately on valid and reliable Numeric Rating Scales.[32,33]

Statistical Analysis

A multistage analysis was employed. All potential effect modifiers were assessed for multicollinearity, which was considered likely if correlations between factors were greater than 0.8.[34] In stage 1, linear mixed models were used to calculate interaction terms as a measure of treatment effect modification. Continuous variables were modeled on their original scale to maximize power and avoid selection of

TABLE 1. Hypothesized Treatment Effect Modifiers

Potential Effect Modifier	Measurement Details	Available Range	Hypothesized Direction of Effect Modification
Örebro Musculoskeletal Pain Screening Questionnaire	A 24-item questionnaire developed to identify participants with low back disorders at risk of developing chronic pain due to psychosocial risk factors, with 21 items rated on a 0–10 scale.[21] Higher scores on the Örebro predict a poorer prognosis in participants with spinal pain.[21,22] The questionnaire has good test-retest reliability.[23]	0–210	Higher scores (indicative of more complex risk factors) favor larger benefits from individualized physical therapy relative to advice.
Coping	The effectiveness of each participant's coping strategies for decreasing pain was assessed using a relevant item from the Örebro Musculoskeletal Pain Questionnaire: *''based on all the things you do to cope, or deal with your pain, on an average day, how much are you able to decrease it''?* This item of the Örebro has been shown to be reliable within the Coping Strategy Questionnaire.[24] Lack of, or ineffective, coping strategies have been shown to be a predictor of poor outcome.[25]	0–10	Higher scores (*i.e.*, poorer coping ability) favor larger benefits from individualized physical therapy relative to advice.
Duration of symptoms	Open question to participants on their baseline questionnaire, which was qualified by the therapist during baseline assessment. The current episode of pain was distinguished from any previous episodes by a 4-week pain-free period.[26]	6–26 wk	Longer duration of symptoms favors larger benefits from individualized physical therapy relative to advice.
Clinical features of inflammation	Four clinical features indicative of inflammation were obtained *via* questionnaire (constant symptoms, waking at night due to pain, morning pain/stiffness >60 min, movement eases symptoms). Preliminary validation of the chosen features has been established by a survey of practitioners regarding indicators of inflammatory low back pain, as well as concurrent validity studies in other inflammatory populations.[27,28]	0–4	Higher number of inflammatory features favors larger benefits from individualized physical therapy relative to advice.
Subgroup membership based on the STOPS classification system	The STOPS classification system divides participants into subgroups based on their hypothesized diagnosis and psychosocial risk.[12] We were interested in whether any of the subgroups derived greater treatment effects than others.	Category:	Larger effect of individualized physical therapy relative to advice in the subgroups with more severe pathoanatomical (DHR subgroup) and psychosocial barriers (MFP subgroup).
		DHR	
		NRDP	
		RDP	
		ZJD	
		MFP	
Proportion of subgroup features above the minimum threshold on the STOPS classification system	We analyzed whether a higher number of subgroup features (indicative of a more definitive classification) predicted response to treatment.	0%–100% above threshold	Larger effects of individualized physical therapy relative to advice in participants with a clearer classification (higher percentage of subgroup features above the minimum threshold).

TABLE 1 (*Continued*)

Potential Effect Modifier	Measurement Details	Available Range	Hypothesized Direction of Effect Modification
Therapist experience	The number of years of experience was recorded for each physical therapist on a continuous scale.	0–17 yr	Given the complexity of the individualized physical therapy treatment, we hypothesized that more experienced physical therapists would achieve a larger treatment effect (individualized physical therapy relative to advice).
Baseline scores on the three primary outcomes (Oswestry, back pain, leg pain)	Activity limitation was assessed using the Oswestry Disability Index,[29] which has been shown to be valid and reliable for use in low back disorder populations.[30,31] Back pain and leg pain intensity were measured separately on valid and reliable Numeric Rating Scales (NRS).[32,33]	0–10 (Back pain, leg pain)	Higher back pain, higher leg pain, and higher activity limitation favors larger benefits from individualized physical therapy relative to advice.
		0%–100% (Activity limitation)	

DHR indicates disc herniation with associated radiculopathy; MFP, multifactorial persistent pain; NRDP, nonreducible discogenic pain; RDP, reducible discogenic pain; STOPS, Specific Treatment of Problems of the Spine; ZDJ, zygapophyseal joint dysfunction.

arbitrary thresholds.[14,35] Models aimed to predict outcome over time generally, where time was included as a repeated measure. Each potential effect modifier was modeled on its own (with variables for time, the predictor, treatment group, and the predictor × treatment interaction term, adjusted for baseline score), with a separate model for each of the three outcomes (back pain, leg pain, and Oswestry). Significant univariate effect modifiers ($P < 0.05$) then progressed to stage 2, where a multivariate analysis was implemented. This commenced with a linear mixed model containing all significant factors from the univariate analysis, with stepwise backwards deletion removing the interaction term with the lowest P value until only significant ($P < 0.05$) terms remained in the final model.[36,37] Missing outcome data were handled *via* restricted maximum likelihood estimation within the linear mixed models,[38] consistent with an intention to treat analysis.

Additional analysis was performed for continuous effect modifiers that attained significance. This involved calculation and graphing of main treatment effects (difference between individualized physical therapy and advice) for various levels of the baseline predictor.[3] This allowed identification of important thresholds for each effect modifier, such as the threshold below which the main treatment effect began to favor advice over individualized physical therapy (if any).

It has been suggested that sample sizes for effect modifier studies need to be four times the size of the primary trial to detect a modifying effect of at least the same magnitude as the overall effect.[7] The time and resources associated with our trial only allowed the sample size to be doubled[11]; hence, the present study is powered to detect treatment modification effects that are approximately 1.2 to 1.4 times the size of the outcomes originally reported in the main trial.[7]

All analyses were undertaken using SPSS-21 and Microsoft Excel.

RESULTS

The results of the RCT have been reported previously.[13] The baseline characteristics of the 300 participants involved in the trial are presented in Table 2, and appeared similar between groups aside from 83% of the individualized physical therapy group being in paid employment at baseline compared to 70% of the advice group. The advice group received significantly more nonmedical cointerventions during the 12-month follow-up period, particularly chiropractic/osteopathy, massage, and group exercise classes.[13,39] Complete baseline data relating to all potential effect modifiers were available for all participants. Outcome data were available for 92% to 97% of participants at each follow-up.[13]

Table 3 shows the interaction terms for each of the effect modifiers included in the univariate analysis. Full model details are presented in the Supplemental Digital Content, http://links.lww.com/BRS/B265. These data show that four statistically significant effect modifiers were identified for at least one outcome measure in the univariate analysis. Higher levels of baseline back pain, higher Örebro scores (indicative of higher risk of persistent pain), longer duration of symptoms, and coping scores indicative of worse coping were associated with larger treatment effects favoring individualized physical therapy over advice.

The results of the final multivariate analysis are presented in Table 4. The final model for each outcome contained two significant effect modifiers: longer duration of symptoms and higher Örebro score for the outcomes of activity limitation and leg pain, and longer duration of symptoms and higher levels of baseline back pain for the back pain outcome. Worse coping did not remain a statistically significant effect modifier in the multivariate models. The direction of all effects was consistent, with higher baseline scores being associated with larger benefits from individualized physical

TABLE 2. Participant Demographics and Baseline Scores on Potential Effect Modifiers

Characteristic	Individualized Physical Therapy (n = 156)	Advice (n = 144)
Demographics		
Age, mean (SD) years	42.7 (11.9)	45.7 (12.3)
Female, number (%)	76 (49%)	71 (49%)
Currently in paid employment, number (%)	129 (83%)	101 (70%)
Baseline scores on potential effect modifiers		
Subgroup membership, number (%)		
Disc herniation with associated radiculopathy	28 (18%)	26 (18%)
Reducible discogenic pain	40 (26%)	38 (26%)
Nonreducible discogenic pain	50 (32%)	46 (32%)
Zygapophyseal joint dysfunction	33 (21%)	31 (22%)
Multifactorial persistent pain	5 (3%)	3 (2%)
Duration of primary pain, mean (SD) weeks	15.4 (6.4)	14.1 (6.1)
Örebro Musculoskeletal Pain Score, mean (SD)	96.4 (24.2)	99.1 (24.9)
Oswestry, mean (SD)	29.2 (11.7)	29.6 (13.0)
Back pain intensity on NRS, mean (SD)	5.3 (2.0)	5.4 (1.9)
Leg pain intensity on NRS, mean (SD)	4.6 (2.8)	4.7 (2.6)
Coping score/10, mean (SD)	5.2 (2.4)	5.0 (2.1)
Therapist experience, mean (SD) years	7.1 (3.8)	6.7 (3.8)
Number of inflammatory features/4, mean (SD)	1.8 (1.0)	1.7 (1.0)
Proportion of subgroup features above threshold, mean (SD) %	35.6 (34.2)	36.7 (37.8)

Oswestry indicates Oswestry Disability Index (10-item questionnaire scored out of 100%); NRS, Numeric Rating Scale (scored from 0 to 10); SD, standard deviation.

therapy over advice. These effect modifiers accounted for between 23% and 34% of the variance in the observed outcome scores in the multivariate models.

The primary univariate models were used to determine baseline score thresholds below which advice was predicted to be superior to individualized physical therapy (Supplemental Digital Content, http://links.lww.com/BRS/B265). These data show that the majority of participants in the trial obtained greater benefit from individualized physical therapy over advice regardless of their baseline scores on any individual effect modifier variable. Very low baseline Örebro scores were, however, associated with greater benefits from advice over individualized physical therapy, applicable to 12.3% of participants with Örebro scores below 70/210 for the activity limitation outcome, or 5.3% of participants with Örebro scores below 62.9 for leg pain outcomes. Baseline back pain scores less than 2.2/10 (activity limitation and leg pain outcomes) or less than 1.2/10 (back pain outcome), along with coping scores less than 0.8/10 (for leg pain outcome) were the only other significant univariate effect modifiers with thresholds below which advice would be superior to individualized physical therapy, but baseline values below these thresholds were present in only 1.0% to 7.4% of the participants involved in the trial.

DISCUSSION

The present study supports the hypothesis that participants with more severe, persistent, or complex low back disorders derive the largest benefits from individualized physical therapy relative to advice. All significant effects were in the direction of our hypothesis. Our modeling suggests that the majority of participants in the trial derived greater benefit from individualized physical therapy over advice unless they scored at the lowest end of the spectrum on multiple baseline predictors evaluated in the present study. The relative benefit of individualized physical therapy over advice increased in participants with higher levels of baseline back pain, longer duration of symptoms, and larger Örebro scores (indicative of higher risk of persistent pain). Higher coping scores (indicating poorer coping) also resulted in larger benefits from individualized physical therapy over advice when considered individually, but this finding did not remain significant in the multivariate analysis. These findings are important because the presence of high scores on these characteristics is typically associated with a worse prognosis and higher treatment and societal costs.[40–42]

The magnitude of effect modification appears clinically important, especially when considering the additive effect of

TABLE 3. Univariate Results of Treatment Effect Modifier Analyses for Each Primary Outcome

Baseline Predictor Variable	Oswestry Outcome β (95% CI), *P*	Back Pain Otcome β (95% CI), *P*	Leg Pain Outcome β (95% CI), *P*
Back pain (0–10) × treatment group (individualized physical therapy) at baseline	−1.37 (−2.58 to −0.15), *P* = 0.027	−0.23 (−0.43 to −0.04), *P* = 0.017	−0.27 (−0.48 to −0.05), *P* = 0.014
Leg pain (0–10) × treatment group (individualized physical therapy) at baseline	−0.72 (−1.52–0.09), *P* = 0.081	−0.04 (−0.16–0.09), *P* = 0.566	−0.07 (−0.23–0.09), *P* = .412
Oswestry (0–10) × treatment group (individualized physical therapy) at baseline	−0.12 (−0.31–0.07), *P* = 0.204	−0.01 (−0.04–0.02), *P* = .580	−0.02 (−0.05–0.02), *P* = 0.262
Örebro score (0–210) × treatment group (individualized physical therapy)	−0.14 (−0.23 to −0.05), *P* = 0.003	−0.01 (−0.03–0.00), *P* = 0.090	−0.02 (−0.04 to −0.01), *P* = 0.009
Duration of primary symptoms (6–26 weeks)* × treatment group (individualized physical therapy)	−0.46 (−0.83 to −0.08), *P* = 0.018	−0.09 (−0.15 to −0.03), *P* = 0.004	−0.08 (−0.15 to −0.01), *P* = 0.023
Coping item of Örebro (0–10)† × treatment group (individualized physical therapy)	−0.75 (−1.80–0.30), *P* = 0.163	−0.10 (−0.26–0.07), *P* = 0.266	−0.20 (−0.39 to −0.00), *P* = 0.046
Number of inflammatory features (0–4) × treatment group (individualized physical therapy)	−0.91 (−3.38–1.55), *P* = 0.467	0.23 (−0.16–0.63), *P* = 0.244	−0.04 (−0.49–0.40), *P* = 0.846
Years of therapist experience (1–13 yr) × treatment group (individualized physical therapy)	−0.38 (−1.01–0.24), *P* = 0.226	0.05 (−0.05–0.15), *P* = 0.336	−0.06 (−0.18–0.05), *P* = 0.294
Proportion of subgroup classification features above the minimum threshold (0%–100%) × treatment group (individualized physical therapy)	0.02 (−0.05–0.08), *P* = 0.626	0.01 (−0.00–0.02), *P* = 0.135	0.01 (−0.00–0.02), *P* = 0.247
Subgroup membership × treatment group (individualized physical therapy)			
NRDP (reference category)	—	—	—
RDP	1.28 (−5.03–7.58), *P* = 0.690	−0.22 (−1.20–0.76), *P* = 0.660	−0.00 (−1.15–1.14), *P* = 0.994
DHR	−1.93 (−8.89–5.04), *P* = 0.586	−0.20 (−1.29–0.89), *P* = 0.718	−0.01 (−1.22–1.21), *P* = 0.992
ZJD	−1.33 (−7.99–5.34), *P* = 0.695	−0.45 (−1.49–0.59), *P* = 0.396	0.09 (−1.16–1.34), *P* = 0.890
MFP	2.16 (−13.25–17.57), *P* = 0.783	1.09 (−1.32–3.50), *P* = 0.375	1.98 (−0.99–4.95), *P* = 0.191

β = Coefficient of the interaction term (predictor × treatment group) from the linear mixed model fitted with the predictor, treatment group, predictor × treatment group interaction, adjustment for baseline scores on the outcome of interest, and time modeled as a repeated measure. Full model parameters are presented in the Supplemental Digital Content, http://links.lww.com/BRS/B265.

Interpretation of effects: the interaction coefficients show the impact of different levels of the predictor variable on the treatment effect (individualized physical therapy relative to advice). Negative coefficients represent a lower outcome score (better outcome) in the individualized physical therapy group relative to the advice group in relation to the predictor variable. For example, for every one point increase in the Örebro (continuous variable) at baseline, the individualized physical therapy group would achieve an additional 0.14-point reduction (improvement) in Oswestry outcome over advice. The reference value for the only categorical variable (subgroup classification) was the NRDP subgroup; hence, a coefficient of −1.93 for the DHR subgroup indicates that individualized physical therapy participants in the DHR subgroup improved (reduced) on the Oswestry by 1.93 additional points (relative to advice) compared to those in the NRDP subgroup.

**Duration of BACK symptoms for all subgroups except for disc herniation with associated radiculopathy subgroup, which was duration of LEG symptoms.*

†For the coping item, the transformed score was used so that higher scores represent poorer coping.

CI indicates confidence interval; DHR, disc herniation with associated radiculopathy; MFP, multifactorial persistent pain; NRDP, nonreducible discogenic pain; RDP, reducible discogenic pain; ZDJ, zygapophyseal joint dysfunction.

TABLE 4. Significant Effect modifiers in the Final Multivariate Models

Baseline Predictor Variable	Oswestry Outcome β (95% CI), *P*	Back Pain Outcome β (95% CI), *P*	Leg Pain Outcome β (95% CI), *P*
Intercept	−8.92 (−16.35 to −1.49), $P=0.019$	0.05 (−0.97–1.06), $P=.930$	−1.60 (−2.95 to −0.24), $P=0.021$
Baseline score of outcome	0.42 (0.31–0.54), $P<0.001$	See below	0.38 (0.29–0.46), $P<0.001$
Time			
5 wk (reference)	—	—	—
10 wk	−3.76 (−5.06 to −2.46), $P<0.001$	−0.38 (−0.62 to −0.14), $P=0.002$	−0.54 (−0.82 to −0.26), $P<0.001$
26 wk	−5.70 (−7.00 to −4.39), $P<0.001$	−0.51 (−0.75 to −0.27), $P<0.001$	−0.70 (−0.98 to −0.42), $P<0.001$
52 wk	−7.95 (−9.26 to −6.63), $P<0.001$	−0.79 (−1.03 to −0.54), $P<0.001$	−0.99 (−1.27 to −0.70), $P<0.001$
Treatment group (individualized physical therapy)	14.84 (4.12–25.56), $P=0.007$	1.54 (0.13–2.94), $P=0.032$	−0.07 (−0.14 to −0.00), $P=0.042$
Baseline back pain (0–10)	—	0.55 (0.41–0.69), $P<0.001$	—
Örebro score (0–210)	0.17 (0.09–0.25), $P<0.001$	—	0.02 (0.01–0.04), $P<0.001$
Duration of primary symptoms (6–26 wk)	0.33 (0.06–0.60), $P=0.018$	0.07 (0.03–0.11), $P=0.002$	0.06 (0.01–0.11), $P=0.015$
Baseline back pain (0–10) × treatment group (individualized physical therapy)	—	−0.23 (−0.42 to −0.04), $P=0.018$	—
Örebro score (0–210) × treatment group (individualized physical therapy)	−0.13 (−0.22 to −0.04), $P=0.005$	—	−0.02 (−0.04 to −0.00), $P=0.013$
Duration of primary symptoms (6–26 weeks) × treatment group (individualized physical therapy)	−0.41 (−0.77 to −0.04), $P=0.030$	−0.09 (−0.15 to −0.03), $P=0.004$	−0.07 (−0.14 to −0.00), $P=0.042$
Estimate of the proportion of variance explained by the model (adjusted-R^2)	34%	23%	29%

Predictor × treatment group interaction term is the degree of effect modification attributable to that predictor in the individualized physical therapy group relative to the advice group (with negative values representing a greater degree of additional improvement in the individualized physical therapy group relative to the advice group).

As a multivariate model, a participant's score on each coefficient can be added to determine their predicted outcome in either the individualized physical therapy or advice group (coefficients with individualized physical therapy in brackets apply only to the individualized physical therapy group).

CI indicates confidence interval.

multivariate models. For example, the main trial found an overall treatment effect (favoring individualized physical therapy over advice) on the Oswestry Disability Scale at 6 months of 5.4 (95% confidence interval: 2.6–8.2).[13] The multivariate results from the effect modifier study, however, show that the expected treatment effect on the Oswestry would vary significantly depending on a participant's baseline Örebro score and duration of symptoms. A person with an average baseline Oswestry score of 29/100, an Örebro score of 130/210 and a duration of symptoms of 24 weeks would be expected to achieve a large clinically significant treatment effect favoring individualized physical therapy over advice (between-group difference of 11.9/100 on the Oswestry). Another person with the same baseline Oswestry score, but an Örebro score of 80/210 and a duration of symptoms of 10 weeks would be expected to achieve a very small treatment effect favoring individualized physical therapy over advice (between-group difference of 0.3/100 on the Oswestry). Such examples suggest that knowing the effect of a participant's baseline Örebro score, back pain score, and duration of symptoms on their likely treatment outcome may be equally or more important than knowing the average treatment effect reported in the primary trial. Other examples can be calculated for any participant using the coefficients in the multivariate models shown in Table 4.

Our study adds to the body of research aiming to identify subgroups of participants who respond more favorably to particular interventions.[4,8,10] The present study appears to

be the first to find longer symptom duration to be an effect modifier for low back disorders, although it is commonly reported as a predictor of poor outcome.[41] One other study has shown Örebro scores to modify treatment effects, with higher scores predicting a better response to motor control *versus* graded activity.[37] Similarly, our study found that higher Örebro scores were associated with greater benefits with an individualized physical therapy intervention that addressed pathoanatomical factors including impaired motor control in addition to psychosocial and neurophysiological factors as part of a biopsychosocial approach. This consistent finding in two studies suggests that physical factors may still be important to consider even in participants with higher psychosocial risk.[43] We are aware of one previous study that found higher back pain scores to be associated with greater effects of a multidisciplinary rehabilitation program relative to usual care,[44] which is consistent with our finding that participants with higher back pain scores achieved better outcomes with a more comprehensive treatment approach (individualized physical therapy) than with advice. The findings of the present study also support the stratified approach to care implemented in the STarT Back Trial, which targeted more comprehensive treatment toward participants with higher risk profiles.[45] Our effect modifier analysis suggests that targeting individualized physical therapy toward participants with greater severity, more persistent symptoms, and higher risk of persistent pain may be warranted.

The strengths of the present study are demonstrated by our compliance with guidelines relating to effect modifier study methods.[5,6,14,15] All potential predictors were selected a priori based on a consistent hypothesis, and were measured in all participants before randomization. Effect modification was assessed using tests of interaction. Multivariate analysis was used to identify independent predictors after adjusting for confounding variables and related predictors. Data relating to all baseline predictors included in the analysis were available for all 300 participants, with outcome data available for more than 90% of participants. Unlike single-group prognostic studies, the use of data from an RCT in this effect modifier study controls for the effect of natural recovery or regression to the mean.[6] These strengths overcome significant methodological weaknesses found in many other low back disorder studies in this field of research.[11]

Despite the strengths of the study, any subgrouping analysis involves a degree of exploration and can only be fully validated once replicated in another population.[6] The findings of the present study were derived from selected participants with noncompensable back pain and/or leg pain of between 6 weeks and 6 months duration, most of whom met the criteria to be classified into three discogenic low back disorder subgroups (76% of the sample) or a zygapophyseal joint dysfunction subgroup (21% of the sample). It is therefore not clear how generalizable the findings would be to broader populations. Although the sample size of our RCT was twice that required to detect primary effects, it may still be underpowered for the purposes of detecting all important treatment effect modifiers.[7,14] The possibility of type I errors is also present, given the assessment of 10 potential effect modifiers across three primary outcomes.[14] We did, however, protect against spurious findings by analyzing prespecified effect modifiers and outcomes on their continuous scales,[14,35] and by employing a multivariate linear mixed model analysis that accounted for correlations between repeated measures across different time points.[46,47]

Advice is recommended in all clinical practice guidelines for low back disorders,[1,48] and our findings support the use of advice in people with short duration, low-intensity back disorders with minimal psychosocial risk. Identifying these baseline characteristics could help healthcare professionals to avoid overtreating people who do not appear to require treatment beyond guideline-based advice. People with low back disorders who present with more complex barriers to recovery (*e.g.*, more persistent symptoms, higher back pain intensity, or presence of psychosocial risk factors), however, appear to require a more comprehensive and individualized treatment approach beyond advice. Studies in this field are important to provide guidance for targeting referrals for individualized physical therapy treatment toward those who are most likely to derive the greatest benefits relative to lower cost and simpler interventions such as advice.

CONCLUSION

People with low back disorders who have higher levels of back pain, higher Örebro scores (indicative of higher risk of persistent pain), and longer duration of symptoms, derive the largest benefit from individualized physical therapy relative to guideline-based advice. People with these characteristics could be preferentially targeted for referral for individualized physical therapy, and for inclusion in future clinical trials. Advice may be sufficient or even preferable in people with very low scores on all of these characteristics.

➢ Key Points

- ❑ People presenting with a low back disorder combined with higher levels of back pain, higher Örebro scores (indicative of higher risk of persistent pain), or longer duration of symptoms derive the largest benefits from individualized physical therapy relative to advice.
- ❑ Those with low scores on all of these features are likely to respond equally well, or potentially better, with advice compared to individualized physical therapy.
- ❑ These findings will help healthcare providers to target individualized physical therapy or advice treatment toward people with low back disorders who are likely to benefit most from one treatment relative to another.

Acknowledgments

The authors wish to acknowledge the trial physical therapists who volunteered to treat participants in this study free of charge. We also acknowledge LifeCare Health for providing facilities, personnel, and resources to allow treatment of participants free of charge.

Trial physical therapists: Aidan Rich, Alexander Chan, Andrew Hahne, Ben Sheat, Christine Baker, Daniel di Mauro, David Goulding, Emma Liu, Gabrielle Hunter, Joel Laing, Justin Moar, Luke Surkitt, Madeleine Ellis, Mark Opar, Matt Richards, Ross Lenssen, Sarah Slater, Shay McLeod

Trial support: Kerryn Dunn

Supplemental digital content is available for this article. Direct URL citations appearing in the printed text are provided in the HTML and PDF version of this article on the journal's Web site (www.spinejournal.com).

References

1. Koes BW, Van Tulder M, Lin CW, et al. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. *Eur Spine J* 2010;19:2075–94.
2. Machado LA, Kamper SJ, Herbert RD, et al. Analgesic effects of treatments for non-specific low back pain: a meta-analysis of placebo-controlled randomized trials. *Rheumatology (Oxford)* 2009;48:520–7.
3. Wang R, Lagakos SW, Ware JH, et al. Statistics in medicine—reporting of subgroup analyses in clinical trials. *N Engl J Med* 2007;357:2189–94.
4. Costa Lda C, Koes BW, Pransky G, et al. Primary care research priorities in low back pain: an update. *Spine (Phila Pa 1976)* 2013;38:148–56.
5. Pincus T, Miles C, Froud R, et al. Methodological criteria for the assessment of moderators in systematic reviews of randomised controlled trials: a consensus study. *BMC Med Res Methodol* 2011;11:14.
6. Hancock M, Herbert RD, Maher CG. A guide to interpretation of studies investigating subgroups of responders to physical therapy interventions. *Phys Ther* 2009;89:698–704.
7. Brookes ST, Whitelya E, Eggerb M, et al. Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. *J Clin Epidemiol* 2004;57:229–36.
8. Gurung T, Ellard DR, Mistry D, et al. Identifying potential moderators for response to treatment in low back pain: a systematic review. *Physiotherapy* 2015;101:243–51.
9. Haskins R, Rivett DA, Osmotherly PG. Clinical prediction rules in the physiotherapy management of low back pain: a systematic review. *Man Ther* 2012;17:9–21.
10. Kent P, Kjaer P. The efficacy of targeted interventions for modifiable psychosocial risk factors of persistent nonspecific low back pain—a systematic review. *Man Ther* 2012;17:385–401.
11. Mistry D, Patel S, Hee SW, et al. Evaluating the quality of subgroup analyses in randomized controlled trials of therapist-delivered interventions for nonspecific low back pain: a systematic review. *Spine (Phila Pa 1976)* 2014;39:618–29.
12. Hahne AJ, Ford JJ, Surkitt LD, et al. Specific treatment of problems of the spine (STOPS): design of a randomised controlled trial comparing specific physiotherapy versus advice for people with subacute low back disorders. *BMC Musculoskelet Disord* 2011;12:104.
13. Ford JJ, Hahne AJ, Surkitt LD, et al. Individualised physiotherapy as an adjunct to guideline-based advice for low back disorders in primary care: a randomised controlled trial. *Br J Sports Med* 2016;50:237–45.
14. Hingorani AD, Windt DA, Riley RD, et al. Prognosis research strategy (PROGRESS): 4 stratified medicine research. *BMJ* 2013;346:e5793.
15. Sun X, Briel M, Walter SD, et al. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *BMJ* 2010;340:c117.
16. Indahl A, Velund L, Reikeraas O. Good prognosis for low back pain when left untampered. A randomized clinical trial. *Spine (Phila Pa 1976)* 1995;20:473–7.
17. Ford JJ, Hahne AJ, Chan AYP, et al. A classification and treatment protocol for low back disorders. Part 3—functional restoration for intervertebral disc related problems. *Phys Ther Rev* 2012;17:55–75.
18. Ford JJ, Richards MC, Hahne AJ. A classification and treatment protocol for low back disorders. Part 4—functional restoration for multi-factorial persistent pain. *Phys Ther Rev* 2012;17:322–34.
19. Ford JJ, Surkitt LD, Hahne AJ. A classification and treatment protocol for low back disorders. Part 2—directional preference management for reducible discogenic pain. *Phys Ther Rev* 2011;16:423–37.
20. Ford JJ, Thompson SL, Hahne AJ. A classification and treatment protocol for low back disorders. Part 1—specific manual therapy. *Phys Ther Rev* 2011;16:168–77.
21. 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;14:209–15.
22. Hockings RL, McAuley JH, Maher CG. A systematic review of the predictive ability of the Orebro Musculoskeletal Pain Questionnaire. *Spine (Phila Pa 1976)* 2008;33:E494–500.
23. Opsommer E, Hilfiker R, Raval-Roland B, et al. Test-retest reliability of the Orebro Musculoskeletal Pain Screening Questionnaire and the Situational Pain Scale in patients with chronic low back pain. *Swiss Med Wkly* 2013;143:w13903.
24. Main CJ, Waddell G. A comparison of cognitive measures in low back pain: statistical structure and clinical validity at initial assessment. *Pain* 1991;46:287–98.
25. Kent PM, Keating JL. Can we predict poor recovery from recent-onset nonspecific low back pain? A systematic review. *Man Ther* 2008;13:12–28.
26. de Vet HC, Heymans MW, Dunn KM, et al. Episodes of low back pain: a proposal for uniform definitions to be used in research. *Spine (Phila Pa 1976)* 2002;27:2409–16.
27. Foldes K, Balint P, Gaal M, et al. Nocturnal pain correlates with effusions in diseased hips. *J Rheumatol* 1992;19:1756–8.
28. Rudwaleit M, Metter A, Listing J, et al. Inflammatory back pain in ankylosing spondylitis: a reassessment of the clinical history for application as classification and diagnostic criteria. *Arthritis Rheum* 2006;54:569–78.
29. Davidson M. Rasch analysis of three versions of the Oswestry Disability Questionnaire. *Man Ther* 2008;13:221–31.
30. Davidson M, Keating JL. A comparison of five low back disability questionnaires: reliability and responsiveness. *Phys Ther* 2002;82:8–24.
31. Frost H, Lamb SE, Stewart-Brown S. Responsiveness of a patient specific outcome measure compared with the Oswestry Disability Index v2.1 and Roland and Morris Disability Questionnaire for patients with subacute and chronic low back pain. *Spine (Phila Pa 1976)* 2008;33:2450–7.
32. Jensen MP, Turner JA, Romano JM, et al. Comparative reliability and validity of chronic pain intensity measures. *Pain* 1999;83:157–62.
33. Pengel LHM, Refshauge KM, Maher CG. Responsiveness of pain, disability, and physical impairment outcomes in patients with low back pain. *Spine (Phila Pa 1976)* 2004;29:879–83.
34. Campbell P, Bishop A, Dunn KM, et al. Conceptual overlap of psychological constructs in low back pain. *Pain* 2013;154:1783–91.
35. Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ* 2006;332:1080.
36. Witt CM, Schutzler L, Ludtke R, et al. Patient characteristics and variation in treatment outcomes: which patients benefit most from acupuncture for chronic pain? *Clin J Pain* 2011;27:550–5.
37. Macedo LG, Maher CG, Hancock M, et al. Predicting response to motor control exercises and graded activity for low back pain patients: preplanned secondary analysis of a randomized controlled trial. *Phys Ther* 2014;94:1543–54.

38. Liu M, Wei L, Zhang J. Review of guidelines and literature for handling missing data in longitudinal clinical trials with a case study. *Pharm Stat* 2006;5:7–18.
39. Hahne AJ, Ford JJ, Surkitt LD, et al. Individualized physical therapy is cost effective compared to guideline-based advice for people with low back disorders. *Spine (Phila Pa 1976)* 2017; 42:E169–76.
40. Dagenais S, Caro J, Haldeman S. A systematic review of low back pain cost of illness studies in the United States and internationally. *Spine J* 2008;8:8–20.
41. Celestin J, Edwards RR, Jamison RN. Pretreatment psychosocial variables as predictors of outcomes following lumbar surgery and spinal cord stimulation: a systematic review and literature synthesis. *Pain Med* 2009;10:639–53.
42. Hockings RL, McAuley JH, Maher CG. A systematic review of the predictive ability of the orebro musculoskeletal pain questionnaire. *Spine (Phila Pa 1976)* 2008;33:E494–500.
43. Ford JJ, Hahne AJ. Pathoanatomy and classification of low back disorders. *Man Ther* 2013;18:165–8.
44. Van der Hulst M, Vollenbroek-Hutten MM, Groothuis-Oudshoorn KG, et al. Multidisciplinary rehabilitation treatment of patients with chronic low back pain: a prognostic model for its outcome. *Clin J Pain* 2008;24:421–30.
45. Hill JC, Whitehurst DG, Lewis M, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet* 2011;378:1560–71.
46. Rihn JA, Hilibrand AS, Radcliff K, et al. Duration of symptoms resulting from lumbar disc herniation: effect on treatment outcomes: analysis of the Spine Patient Outcomes Research Trial (SPORT). *J Bone Joint Surg Am* 2011;93: 1906–14.
47. West BT, Welch KB, Galecki AT. *Linear Mixed Models: A Practical Guide Using Statistical Software ed*. Boca Raton, FL: Chapman & Hall/CRC; 2007.
48. Dagenais S, Tricco AC, Haldeman S. Synthesis of recommendations for the assessment and management of low back pain from recent clinical practice guidelines. *Spine J* 2010;10:514–29.