People with low back pain typically need to feel ‘much better’ to consider intervention worthwhile: an observational study

Manuela Loureiro Ferreira1,2, Paulo Henrique Ferreira1, Robert Dale Herbert3 and Jane Latimer3

1The University of Sydney, Australia, 2Pontificia Universidade Catolica de Minas Gerais, Brazil, 3The George Institute for International Health, Australia

**Question:** How much of an effect do five common physiotherapy interventions need to have for patients with low back pain to perceive they are worth their cost, discomfort, risk, and inconvenience? Are there any differences between the interventions? Do specific characteristics of people with low back pain predict the smallest important difference?

**Design:** Cross-sectional, observational study. **Participants:** 77 patients with non-specific low back pain who had not yet commenced physiotherapy intervention. **Outcome measures:** The smallest worthwhile effect was measured in terms of global perceived change (0 to 4) and percentage perceived change. **Results:** Participants perceived that intervention would have to make them ‘much better’, which corresponded to 1.7 (SD 0.7) on the 4-point scale, or improve their symptoms by 42% (SD 23), to make it worthwhile. There was little distinction made between interventions, regardless of whether smallest worthwhile effects were quantified as global perceived change ($p = 0.09$) or percentage perceived change ($p = 1.00$). Severity of symptoms independently ($p = 0.01$) predicted percentage perceived change explaining 9% of the variance, so that for each increase in severity of symptoms of 1 point out of 10 there was an increase of 4% in the percentage perceived change that participants considered would make intervention worthwhile. **Conclusions:** Typically people with low back pain feel that physiotherapy intervention must reduce their symptoms by 42%, or make them feel ‘much better’ for intervention to be worthwhile. [Ferreira ML, Ferreira PH, Herbert RD, Latimer J (2009) People with low back pain typically need to feel ‘much better’ to consider intervention worthwhile: an observational study. *Australian Journal of Physiotherapy* 55: 123–127]

**Keywords:** Clinical Significance, Low back pain, Data interpretation statistical, Health care surveys, Effect of intervention, Rehabilitation, Physiotherapy

**Introduction**

The main intent of randomised clinical trials and systematic reviews is to provide estimates of the effect of intervention. Interpretation of the statistical significance of the estimated effect of intervention is usually not problematic, but interpretation of clinical significance (or clinical importance) can be difficult (Chan 2001).

Several approaches have been used to investigate the clinical significance of the effect of interventions on health-related quality of life (Bombardier 2001, Cella 2002; Chan 2001, Devereaux 2001, Farrar 2001, Gallagher 2001, Guyatt 1998, Haag 2003, Man-son-Hing 2002, Middel 2001, Norman 2001, Redelmeier 1996, Samsa 1999, Schunemann 2003, van Walraven 1999, Yelland and Schluter 2006, Zisapel 2003). Most studies have assessed clinical significance by determining how large the effect must be for patients to say the intervention made them ‘a bit better’ or ‘much better’. Threshold values (the degree of patient-rated change considered to be clinically important, eg, ‘much better’) are usually nominated by researchers or clinicians (Wells 2001). Such studies do not directly assess how beneficial the intervention must be for patients to feel that the intervention was worth receiving.

Only one study has sought patients’ opinions on what constitutes the minimum worthwhile reduction in symptoms of low back pain. Yelland and Schluter(2006) asked 110 patients with chronic low back pain about both their desired reduction in symptoms as well as the minimum reduction in symptoms they would expect for the intervention to be considered worthwhile. The minimum worthwhile reduction for pain was 25% and for disability was 35%. It was not clear, however, whether patients considered the discomforts, risks, and inconveniences of the intervention when making these decisions (Yelland and Schluter 2006).

In our opinion, the decision of whether an intervention is clinically significant must involve consideration of whether the estimated effect of intervention is big enough to be worth its costs, discomforts, risks, and inconveniences. Barrett and colleagues (Barrett 2005, Barrett 2007) have called this construct the ‘sufficiently important difference’. We will refer, synonymously, to the ‘smallest worthwhile effect’. This construct has three characteristics. First, it can only be evaluated by the beneficiary of care (usually the patient). Second, because this decision involves consideration of the cost, discomfort, risk, and inconvenience of the intervention, the estimate of what constitutes the smallest important difference must generally be intervention-specific. Finally, because the sufficiently important difference is the effect of intervention it must be thought of as the hypothetical difference between the outcome a person would experience if they had the intervention and the outcome the same person would have if they had no intervention.
The aim of this study was to assess patients’ perceptions of what constitutes the smallest worthwhile effect of specific interventions. We sought the opinions of patients with non-specific low back pain about a range of commonly-administered conservative interventions. Our specific research questions were:

1. What is the smallest effect perceived by patients with non-specific low back pain to make five common physiotherapy interventions worth their cost, discomfort, risk and inconvenience?
2. Are there any differences in smallest worthwhile effect between the interventions?
3. Do specific characteristics of people with low back pain (age, duration of symptoms, severity of symptoms, and past experience with intervention) predict the smallest worthwhile effect?

**Method**

**Design**

A cross-sectional study was conducted involving people with non-specific low back pain. Participants were interviewed before commencing physiotherapy intervention at a large hospital outpatient department. By interviewing prior to intervention we avoided contaminating perceptions of smallest important difference with improvement or deterioration in symptoms due to the intervention. Each participant was told about five physiotherapy interventions commonly provided for people with non-specific low back pain (exercise, spinal manipulation, ultrasound, local heat, and massage) (Turner 2002). The interventions were described using a standardised script which outlined how the intervention was administered, the usual number and length of intervention sessions, and the proposed benefits and risks of intervention (see Appendix 1 on eAddenda for script). Thus the patient was familiar with the interventions before he or she was asked about what constituted a worthwhile effect of intervention.

**Participants**

Consecutive patients with non-specific low back pain presenting to an outpatient physiotherapy department in a large teaching hospital were included in the study. Participants were excluded if they were aged less than 18 or more than 80 years, or if they had been diagnosed by the referring medical practitioner as having specific spinal pathology (nerve root involvement, inflammatory disorders, fracture, or malignancy).

**Measurement of smallest important difference**

The smallest worthwhile effect was measured in terms of both global perceived change and percentage perceived change. Global perceived change was measured by asking participants to rate the smallest important difference where 0 = ‘no better’, 1 = ‘a little better’, 2 = ‘much better’, 3 = ‘very much better’ and 4 = ‘fully recovered’. Percentage perceived change was measured by asking participants to rate the smallest worthwhile effect on a visual analogue scale where 0% indicated ‘no better’ and 100% indicated ‘fully recovered’. The same questions were asked in regard to each of the five interventions (Box 1).

**Data analysis**

Responses were summarised with descriptive statistics. ANOVA was performed to investigate differences in response among the five interventions on continuous measures of effect. Friedman’s test was performed to analyse differences in response across interventions of discrete variables. Multiple linear regression was used to predict the smallest worthwhile effect expressed as global perceived change, and percentage perceived change based on four explanatory factors. Predictors included in the model were: past experience with all interventions (total number of sessions across all interventions), severity of symptoms in the past seven days, age, and duration of symptoms in weeks. Predictors were chosen a priori and forced into the model, i.e., a selection procedure was not used. The significance level was set at 0.05.

**Results**

**Participants**

Eighty-eight consecutive patients with low back pain were invited to participate in the study and 77 (88%) agreed to participate. Eleven (12%) patients declined to participate because they were unable to speak English (n = 4), were unable to attend the appointment (n = 3), did not have non-specific low back pain (n = 2), or were not willing to participate (n = 1). Fifty-one (66%) participants were female with a median duration of symptoms of 4 weeks (IQR 9, range 1 week to 40 years). Fifty-five (42%) participants had previously experienced at least one of the five interventions for low back pain, exercise being the most commonly experienced (22%), followed by spinal manipulation (20%), local heat (18%), massage (15%), and ultrasound (8%). Characteristics of the 77 participants are presented in Table 1.

**Table 1. Mean (SD) characteristics of participants.**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>(n = 77)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>53.2 (15.1)</td>
</tr>
<tr>
<td>Severity of symptoms (1 to 10)</td>
<td>6.9 (2.1)</td>
</tr>
<tr>
<td>Past experience with intervention</td>
<td></td>
</tr>
<tr>
<td>(number of sessions for participants</td>
<td></td>
</tr>
<tr>
<td>with past experience)</td>
<td></td>
</tr>
<tr>
<td>Exercise</td>
<td>4.1 (10.0)</td>
</tr>
<tr>
<td>Spinal manipulation</td>
<td>3.0 (7.9)</td>
</tr>
<tr>
<td>Ultrasound</td>
<td>0.9 (3.9)</td>
</tr>
<tr>
<td>Local heat</td>
<td>2.1 (5.3)</td>
</tr>
<tr>
<td>Massage</td>
<td>1.6 (4.2)</td>
</tr>
</tbody>
</table>
What is the smallest effect that would make intervention worthwhile?

Group data of the ‘smallest worthwhile effect’ in terms of global perceived change and percentage perceived change for each intervention is presented in Table 2 while individual data are presented in Table 3 (see eAddenda for Table 3). On average, participants perceived that intervention would have to make them ‘much better’ which corresponded to 1.7 (SD 0.7) on the 4-point scale, or improve their symptoms by 42% (SD 23) to make it worthwhile.

Are there differences in smallest worthwhile effects between the interventions?

Table 2 shows that participants made little distinction between interventions, regardless of whether smallest worthwhile effect were quantified as global perceived change ($p=0.09$) or percentage perceived change ($p=1.00$). For global perceived change, agreement was exact in 76% of pairs across the five interventions. For percentage perceived change, the ICC(2,1) was 0.95 (95% CI 0.94 to 0.97, $p<0.001$) across the five interventions.

Do characteristics of people with low back pain predict the smallest worthwhile effect?

Since no systematic difference was observed across interventions, the remaining analyses were performed using the mean of each participant’s responses for the five interventions. Severity of symptoms independently ($p=0.01$) predicted percentage perceived change, explaining 9% of the variance. For each increase of 1 point out of 10 in severity of symptoms, there was an increase of 4% in the percentage perceived change that participants considered would make intervention worthwhile (Table 4). Age, duration of symptoms, or past experience with any of the five interventions did not predict either global or percentage perceived change, regardless of whether past experience was expressed as number of sessions or dichotomously as YES/NO.

### Discussion

The aim of this study was to determine what patients with low back pain consider to be a worthwhile effect of intervention. In past clinical trials of intervention for low back pain, decisions about the smallest important difference (‘clinical significance’) have mostly been derived from distribution-based and anchor-based studies. These approaches rely either on clinicians’ perspectives or on the property of the measurement instrument to establish a worthwhile effect of intervention (Wells 2001) but fail to consider patients’ opinions. In contrast, our study used a patient-centred approach and is therefore more likely to reflect accurately what patients perceive to be clinically worthwhile. Moreover, our patients’ perceptions involved the consideration of costs, risks, and inconveniences of intervention before the commencement of intervention, providing a reference by which intervention success can be prospectively measured.

We showed that, on average, participants need to experience a 42% improvement in symptoms to consider physiotherapy intervention worthwhile. This corresponds to patients considering themselves ‘much better’ as a result of intervention. Yelland and Schluter (2006) found that a median reduction in pain of 25% and in disability of 35% is necessary for patients with chronic low back pain to
interventions. For instance, of randomised trials listed in trials are designed to investigate the efficacy of individual of combined interventions because most randomised clinical trials and systematic reviews, then it should not worthwhile effect is to aid in interpreting randomised for all combinations. Moreover, if determining the smallest worthwhile effect would be difficult to evaluate the smallest worthwhile effect of the combined interventions. However, there are numerous be important to evaluate the smallest important difference interventions separately. We acknowledge that it would also participants were asked to consider each of the five When determining the smallest worthwhile effect, patients were asked to consider each of the five interventions separately. We acknowledge that it would also be important to evaluate the smallest important difference of the combined interventions. However, there are numerous possible combinations of interventions for back pain, and it would be difficult to evaluate the smallest worthwhile effect for all combinations. Moreover, if determining the smallest worthwhile effect is to aid in interpreting randomised clinical trials and systematic reviews, then it should not be necessary to investigate the smallest worthwhile effect of combined interventions because most randomised trials are designed to investigate the efficacy of individual interventions. For instance, of randomised trials listed in PEDro over the last 10 years, 79% of the trials on massage, 50% of the trials on heat and ultrasound, 60% of the trials on exercise, and two-thirds of those on spinal manipulation investigate the efficacy of these interventions in isolation.

The only significant predictor of the size of patients’ perceptions of a worthwhile reduction in symptoms in the present study was initial symptom severity. Age, duration of symptoms, and past experience with interventions did not appear to influence them. Kelly reported similar results in an acute emergency department when calculating an anchor-based, clinically-significant difference in pain (Kelly 1998) and concluded that 9% was the minimum difference that patients were able to perceive as a reduction in pain, and that this was not associated with age, gender, cause of pain, or pain severity. Likewise, Yelland and Schluter (2006) found no associations between minimum clinically-worthwhile changes in low back pain and age, pain duration, pain intensity, depression, or anxiety.

The smallest worthwhile effect is an effect of intervention (Ferreira and Herbert 2008), ie, it refers to the hypothetical difference between the outcome a patient would experience with and without intervention. While we attempted to communicate this complex idea to participants, we believe most patients could not make this distinction in our study. For that reason, the estimates presented in our study represent measures of symptom reduction from baseline and should be used with care when interpreting the difference in outcome between two interventions.

In conclusion, typically people with non-specific low back pain feel that physiotherapy intervention must reduce their symptoms by nearly half and make them feel ‘much better’ if intervention is to be worthwhile.

Table 4. Regression coefficients (95% CI) for each predictor of global perceived change and percentage perceived change.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Global perceived change (n = 76)</th>
<th>Percentage perceived change (n = 77)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>−0.01 (−0.02 to 0.00)</td>
<td>−0.32 (−0.70 to 0.07)</td>
</tr>
<tr>
<td>Duration of symptoms</td>
<td>−0.00 (−0.02 to 0.01)</td>
<td>0.05 (−0.01 to 0.02)</td>
</tr>
<tr>
<td>Severity of symptoms</td>
<td>0.05 (−0.02 to 0.11)</td>
<td>3.66 (0.90 to 6.38)</td>
</tr>
<tr>
<td>Past experience with intervention</td>
<td>0.00 (−0.01 to 0.00)</td>
<td>−0.19 (−0.46 to 0.08)</td>
</tr>
<tr>
<td>Constant</td>
<td>1.66 (1.04 to 2.29)</td>
<td>34.06 (8.02 to 60.12)</td>
</tr>
</tbody>
</table>

consider prolotherapy worthwhile (Yelland and Schluter 2006). The larger smallest worthwhile effect found in our study may be due to our patients being asked about general symptom reduction rather than specific pain or disability. This suggests that patients have different requirements and expectations of intervention according to the nature of the outcome.

When our results are compared to those of anchor-based studies, a marked difference can be seen. Anchor-based studies suggest that a reduction in disability of 10% (Haag 2003, Kelly 2001), and a decrease in pain of 18% (Haag 2003, Farrar 2001) represent clinically-important effects. Patients’ requirements might be somewhat underestimated if these values are accepted, since a 42% reduction of initial symptoms corresponds to an absolute symptom reduction of 3 points out of 10 in our study.

There was a large variability in responses by participants in the present study as evidenced by the large standard deviations. The high level of variability in perceptions of the smallest worthwhile effect may be due to the heterogeneity in symptoms in terms of duration and severity. High levels of variability between patients have also been observed in other studies of patients’ expectations of intervention (Barrett 2005, Barrett 2007, Yelland and Schluter 2006). For example, Yelland and Schlutter’s results for pain show a high degree of variability, with 50% of their sample requiring at least 20% improvement of pain, and a quarter requiring at least 50% improvement to consider intervention worthwhile.

Patients made little distinction between interventions when considering what would represent a worthwhile reduction in symptoms. It is interesting to consider what factors may have influenced this decision. All participants were read a script about the cost, discomfort, risk, and inconvenience of the interventions which suggests that the consistency in the smallest worthwhile effect across interventions was dictated by factors that were similar in all interventions, eg, the number of times they had to visit a health practitioner.

When determining the smallest worthwhile effect, patients were asked to consider each of the five interventions separately. We acknowledge that it would also be important to evaluate the smallest important difference of the combined interventions. However, there are numerous possible combinations of interventions for back pain, and it would be difficult to evaluate the smallest worthwhile effect for all combinations. Moreover, if determining the smallest worthwhile effect is to aid in interpreting randomised clinical trials and systematic reviews, then it should not be necessary to investigate the smallest worthwhile effect of combined interventions because most randomised trials are designed to investigate the efficacy of individual interventions. For instance, of randomised trials listed in

eAddenda: Appendix 1, Table 3 available at AJP.physiotherapy.asn.au

Ethics: Ethics Committees of The University of Sydney and the South Western and Western Sydney Area Health Services approved this study. Informed consent was gained from all participants before data collection began.

Support: Rob Herbert is supported by the National Health and Medical Research Council of Australia. Manuela Ferreira is a post-doctoral research fellow supported by the Faculty of Health Sciences, The University of Sydney.

Acknowledgements: We gratefully acknowledge Matthew Jennings for his assistance with patient recruitment.

Correspondence: Manuela L Ferreira, Clinical & Rehabilitation Sciences Research Group, The University of Sydney, PO BOX 170, Lidcombe NSW 1825 Australia. Email: mferreira@usyd.edu.au
References


Website
www.pedro.org.au