



## Research

## The smallest worthwhile effect on pain intensity of nonsteroidal anti-inflammatory drugs and exercise therapy for acute and chronic low back pain: a benefit-harm trade-off study

Harrison J Hansford <sup>a,b</sup>, Matthew D Jones <sup>a,b</sup>, Aidan G Cashin <sup>a,b</sup>, Raymond WJG Ostelo <sup>c,d</sup>, Alessandro Chiarotto <sup>e</sup>, Sam A Williams <sup>b</sup>, Saurab Sharma <sup>a,b</sup>, Jack J Devonshire <sup>b</sup>, Michael C Ferraro <sup>a,b</sup>, Michael A Wewege <sup>a,b</sup>, James H McAuley <sup>a,b</sup>

<sup>a</sup> School of Health Sciences, University of New South Wales, Sydney, Australia; <sup>b</sup> Centre for Pain IMPACT, Neuroscience Research Australia, Sydney, Australia; <sup>c</sup> Department of Health Sciences, Faculty of Science and Amsterdam Movement Science Research Institute, Vrije Universiteit, Amsterdam, The Netherlands; <sup>d</sup> Department of Epidemiology and Data Science, Amsterdam University Medical Centre, Vrije Universiteit, Amsterdam, The Netherlands; <sup>e</sup> Department of General Practice, Erasmus MC, University Medical Centre, Rotterdam, The Netherlands

## KEY WORDS

Clinical importance  
Patient-reported outcome measures  
Smallest worthwhile effect  
Low back pain  
Exercise

## ABSTRACT

**Question:** What are the smallest worthwhile effects of nonsteroidal anti-inflammatory drugs (NSAIDs) for people with acute and chronic low back pain (LBP)? What is the smallest worthwhile effect of individualised exercise for people with chronic LBP compared with no intervention? **Design:** Benefit-harm trade-off study. **Participants:** Participants were recruited by advertisement on social media and included if they were English-speaking adults in Australia who had non-specific LBP. **Outcome measure:** Pain intensity. **Results:** A total of 116 people with acute LBP and 230 people with chronic LBP were recruited. For acute LBP, the smallest worthwhile effect of NSAIDs additional to no intervention was a 30% (IQR 10 to 40%) reduction in pain intensity. For chronic LBP, the smallest worthwhile effect of NSAIDs additional to no intervention was a 27.5% (IQR 10 to 50%) reduction in pain intensity. For chronic LBP, the smallest worthwhile effect of exercise additional to no intervention was a 20% (IQR 10 to 40%) reduction in pain intensity. There were small associations between baseline pain, duration of pain and level of exercise and the smallest worthwhile effect of NSAIDs for acute LBP. There were no other clear associations. **Conclusions:** For people with LBP, the smallest worthwhile effect of exercise and NSAIDs additional to no intervention is approximately a 20 to 30% reduction in pain. These results can inform the interpretation of the effects of NSAIDs and exercise in randomised trials and meta-analyses, incorporating consumers' perspectives. Further research on comparisons between different interventions and on other core LBP outcomes may inform decision-making. **Registration:** OSF [osf.io/3erjx/](https://osf.io/3erjx/). [Hansford HJ, Jones MD, Cashin AG, Ostelo RWJG, Chiarotto A, Williams SA, Sharma S, Devonshire JJ, Ferraro MC, Wewege MA, McAuley JH (2023) The smallest worthwhile effect on pain intensity of nonsteroidal anti-inflammatory drugs and exercise therapy for acute and chronic low back pain: a benefit-harm trade-off study. *Journal of Physiotherapy* ■-■-■]

© 2023 Published by Elsevier B.V. on behalf of Australian Physiotherapy Association. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Introduction

Low back pain (LBP) has been the leading cause of years lived with disability globally for 3 decades, with an estimated 568 million people reporting LBP in 2019.<sup>1</sup> It has a lifetime prevalence of up to 80% and half of those people seek care for their LBP.<sup>2</sup> For 80 to 90% of patients a specific patho-mechanical cause cannot be reliably identified and LBP is considered to be non-specific.<sup>3</sup>

To manage non-specific LBP, clinical practice guidelines commonly recommend nonsteroidal anti-inflammatory drugs (NSAIDs) for acute/subacute LBP (< 3 months) and chronic LBP (> 3 months), as well as exercise for chronic LBP.<sup>4</sup> For people with LBP, pain is a common reason for seeking care<sup>5</sup> and pain relief is an important outcome of

intervention.<sup>6-8</sup> Systematic reviews show that, compared with placebo, NSAIDs reduce pain intensity on a 0-to-10 numerical pain rating scale (NPRS) by 0.7 (95% CI 0.4 to 1.1) points in acute LBP<sup>9</sup> and by 0.7 (95% CI 0.3 to 1.1) points in chronic LBP.<sup>10</sup> For chronic LBP, exercise compared with no intervention or usual care reduces pain intensity by 1.5 (95% CI 1.2 to 1.8) points on a 0-to-10 NPRS.<sup>11</sup> It is unclear whether people with LBP consider effects of this magnitude worthwhile.<sup>12</sup>

The minimum benefit of an intervention that patients consider worthwhile given the costs, risks and inconveniences is often called the smallest worthwhile effect.<sup>13-15</sup> Estimates of the smallest worthwhile effect can inform research, clinical and policy decisions. Benefit-harm trade-off is a frequently used method used to estimate the smallest worthwhile effect.<sup>14,16-21</sup> Benefit-harm trade-off provides the

patients' perspectives on the smallest difference between interventions that they would consider to be worth undertaking the intervention compared with an alternative intervention.<sup>12,15,22–24</sup> A discrete choice experiment is another method used to estimate the smallest worthwhile effect; however, there have previously been discrepancies between estimates of the smallest worthwhile effect from benefit-harm trade-off and discrete choice experiment methods.<sup>20</sup>

The smallest worthwhile effect for people with chronic LBP has been estimated to be a 20% reduction in pain intensity for physiotherapy,<sup>14,19</sup> and a 30% reduction for NSAIDs,<sup>14</sup> in addition to no intervention (ie, natural history, regression to the mean, etc).<sup>14</sup> The utility of the estimates of the smallest worthwhile effect of physiotherapy is limited, as these interventions do not reflect the more specific and contemporary recommendations for LBP.<sup>4</sup> There are no published estimates of the smallest worthwhile effect of guideline-recommended interventions such as NSAIDs for people with acute LBP and individualised exercise for people with chronic LBP. The smallest worthwhile effect of NSAIDs for chronic LBP has been reported;<sup>14</sup> however, this has not been replicated.

Providing guidance on the smallest worthwhile effect for guideline-recommended interventions will help stakeholders (including researchers, clinicians and policymakers) interpret between-group effects in randomised trials and meta-analyses, and implement and evaluate interventions.

Therefore, the specific research questions for this benefit-harm trade-off study were:

1. What are the smallest worthwhile effects of NSAIDs for people with acute and chronic LBP?
2. What is the smallest worthwhile effect of individualised exercise for people with chronic LBP compared with no intervention?

## Methods

### Design

This was a benefit-harm trade-off study. The protocol was registered on the Open Science Framework<sup>25</sup> prior to data collection. Considering that there is no specific reporting guideline for studies using the benefit-harm trade-off method, the findings are reported following the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guideline.<sup>26</sup> The study conformed to the Declaration of Helsinki. Patients were not involved when designing this study.

### Eligibility criteria

Patients were included if they were aged  $\geq 18$  years, proficient in English and reported having current non-specific LBP, defined as pain lasting for  $\geq 1$  day, in the area on the posterior aspect of the body from the lower margin of the 12<sup>th</sup> ribs to the gluteal folds.<sup>1</sup> Patients were excluded if they had recent surgery ( $< 3$  months ago) to the lower limbs or back, or if they were living outside of Australia (due to differing costs of interventions globally).

### Recruitment

Participants were recruited by targeted advertisements on online social media (Facebook, Twitter and TikTok) between 15 July 2021 and 30 August 2022. Paid advertisements were limited to users of Facebook in Australia aged  $> 18$  years. Other terms of interest were used for targeting the advertisements (Appendix 1 on the eAddenda). Participants were also recruited through email contact of people who expressed interest in participating in future research from our group.

### Data collection

Demographic and descriptive data including sex, previous experience with the intervention/s (NSAIDs and/or exercise), perceived

general health status using a 5-point Likert scale,<sup>27</sup> self-reported duration of weekly physical activity using a 5-point Likert scale, pain intensity over the past 24 hours using the 0-to-10 NPRS<sup>28</sup> and pain duration were collected using an online survey using Research Electronic Data Capture (REDCap).<sup>29,30</sup> The survey codebook is available on OSF (osf.io/3erjx/) and Appendix 2 on the eAddenda.

### Benefit-harm trade-off

Benefit-harm trade-off<sup>14,19,20</sup> was used to estimate the smallest worthwhile effect using an online REDCap<sup>29,30</sup> survey. Participants were presented with a description of the intervention explaining the costs, potential risks and inconveniences (Appendix 3 on the eAddenda). Participants were then asked if they would consider the intervention worthwhile with a complete reduction in pain. If participants would not consider the intervention worthwhile, they did not complete the benefit-harm trade-off questions and were not included in the estimation of the smallest worthwhile effect because there was no minimum benefit these participants would consider worthwhile. Participants who would consider the intervention worthwhile were then asked to provide the percentage of pain reduction that they would require to make the intervention worthwhile that was in addition to no intervention (ie, natural history, regression to the mean, etc) – described to participants as a 30% reduction for acute LBP and a 20% reduction for chronic LBP in the next week<sup>31</sup> – while considering the potential costs, risks and inconveniences (Figure 1). Pain intensity was selected as the outcome of interest, as it is an outcome that is important to patients<sup>32</sup> and is commonly used in reports of randomised trials as a primary outcome.<sup>33</sup>

As guidelines suggest that exercise should be individualised, participants were asked to design their 'ideal' exercise program based on exercise modality (aerobic, Pilates/yoga, strengthening and stabilisation), weekly duration ( $< 60$  minutes, 60 to 90 minutes, 90 to 120 minutes, 120 to 150 minutes or  $> 150$  minutes) and level of supervision (supervised fortnightly by a health professional or not). It was possible to choose multiple modalities of exercise (ie, a participant's ideal program could include both aerobic and strengthening exercise). The 'exercise program' that participants chose was then shown to them, and the respective costs and inconveniences were outlined. Participants provided responses to the benefit-harm trade-off questions based on their individualised program. Complete benefit-harm trade-off questions for all interventions can be found in Appendix 3.

### Data analysis

#### Sample size calculation

There are no guidelines for estimating the sample size required to estimate the smallest worthwhile effect; therefore, it was decided to power the study on the regression analyses. Using an online sample size calculator,<sup>34</sup> it was determined that 182 participants (91 each of acute and chronic LBP) were required to conduct independent linear regressions of six baseline characteristics and the smallest worthwhile effect in the benefit-harm trade-off study with  $\alpha = 0.05$  and power = 0.8. This approach is in line with previous benefit-harm trade-off studies.<sup>14,19</sup>

#### Calculation of the smallest worthwhile effect

Packages<sup>35–42</sup> in R software<sup>a</sup> were used to organise and analyse the data.<sup>43</sup> The smallest worthwhile effect was estimated as the median and interquartile range (IQR) of the lowest reduction in pain intensity (%) that participants considered worthwhile in addition to no intervention. All other results are presented as mean and standard deviation unless otherwise indicated. Univariable linear regression models were used to estimate associations between baseline characteristics (ie, pain intensity, duration of LBP, previous usage of NSAIDs or exercise, sex, perceived health status, level of exercise) and the smallest worthwhile effect of NSAIDs (acute and chronic LBP) and exercise (chronic LBP only). The non-parametric bootstrap with 1,000 replicates was used to estimate 95% confidence intervals (CI) for the observed associations.<sup>44</sup>

<p>If you didn't take the pills your back pain could get about 20% better in the next week as this is the natural course of low back pain.</p>	<input checked="" type="button" value="Yes"/> <input type="button" value="No"/>	<input type="button" value="reset"/>
<p>If you got a complete reduction in pain with the anti-inflammatory pills would you consider them worthwhile?</p>		
<p>If you didn't take the drugs your back pain could get about 20% better in the next week as this is the natural course of low back pain.</p>	<input type="text" value="40"/>	<p>number is in percentage improvement in pain intensity</p>
<p>However, if you were given the anti-inflammatory drugs, how much additional reduction of your pain would you need to see to make the drugs worthwhile?</p>		
<p>Now, suppose you only got 30% better on top of the initial 20% improvement (ie, a 50% total reduction).</p>	<input checked="" type="button" value="Yes"/> <input type="button" value="No"/>	<input type="button" value="reset"/>
<p>Would that be enough of a reduction in pain to make the drugs worth their cost and side-effects?</p>		
<p>Now, suppose you only got 20% better on top of the initial 20% (ie, a 40% total reduction).</p>	<input type="button" value="Yes"/> <input checked="" type="button" value="No"/>	<input type="button" value="reset"/>
<p>Would that be enough of a reduction in pain to make the drugs worth their cost and side-effects?</p>		

**Figure 1.** An example of the benefit-harm trade-off questions for nonsteroidal anti-inflammatory drugs asked to participants; the smallest worthwhile effect for this participant would be 30%.

To assist with interpretation of the results, the percentage of the smallest worthwhile effect was converted to between-group point-estimates representing different hypothetical levels of baseline pain (4, 5, 6, 7, 8 on a 0-to-10 NPRS).<sup>45</sup> These point estimates of the smallest worthwhile effect were calculated as below:

$$\text{smallest worthwhile effect ( / 10)} = \text{smallest worthwhile effect (\%)} \\ \times \text{baseline NPRS pain ( / 10)}$$

### Post-hoc analyses

We aimed to explore the impact that the description of the risk of interventions had on the smallest worthwhile effect. We made no changes to the inclusion criteria, data collection or sample size requirements. The first pre-specified survey described NSAIDs as having rare but serious risks, and exercise having minor risks, similar to previous benefit-harm trade-off studies.<sup>14</sup> The second survey modified the descriptions of risks, with NSAIDs described as having minor risks and exercise having rare but serious risks. Complete descriptions are displayed in Appendix 3. We did not modify the description of NSAIDs for people with acute LBP due to the lack of comparison. A more detailed description of the exploratory investigation can be seen in Appendix 4 on the eAddenda. We report all post-hoc analyses as sensitivity analyses. The second sample was recruited, following the same methods outlined earlier, between 12 May 2022 and 30 August 2022. A Shapiro-Wilk test was used to assess data normality.<sup>46</sup> To assess whether the distribution of the smallest worthwhile effect differed between the two samples recruited, where appropriate we used a Mann-Whitney U test<sup>47</sup> in the presence of non-normally distributed data and an independent *t*-test when data were normally distributed. If assumptions of the statistical tests were violated, qualitative differences were presented in medians/means. Participants who responded to the both the first and second surveys were removed from the second analysis.

## Results

### Compliance with the study protocol

We completed the methods described in the registered protocol and the post-hoc analysis described above.

### Characteristics of study participants

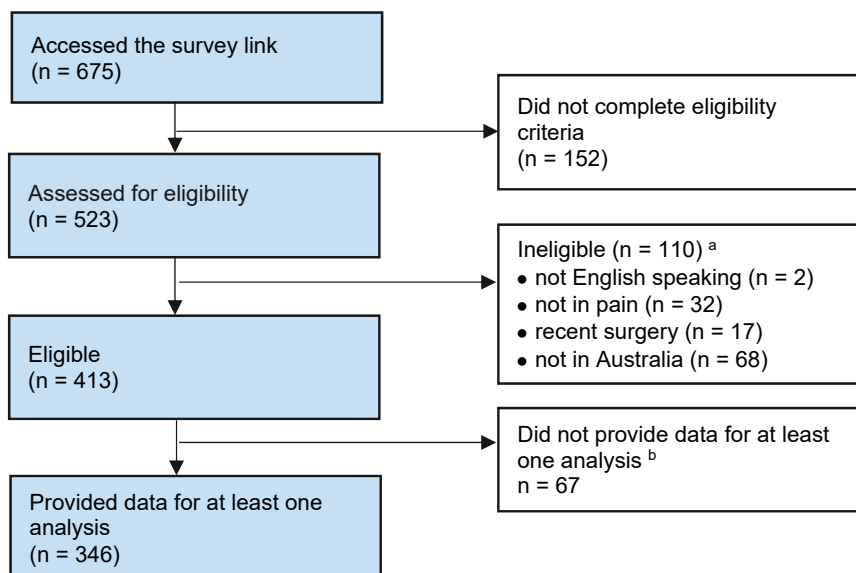
The initial survey link was accessed by 675 people, of whom 413 were eligible. Of those who were eligible, 346 participants (*n* = 116 acute, *n* = 230 chronic) provided data for at least one analysis (84% completion rate) (Figure 2). Compared with completers, participants who did not complete the survey were older and reported higher pain intensity (Appendix 5 on the eAddenda). Characteristics of participants included in the study are presented in Table 1.

### Smallest worthwhile effect of NSAIDs for acute LBP

Twenty (18%) participants reported that they would not consider taking NSAIDs for their LBP. For people with acute LBP who would consider NSAIDs (*n* = 96), the smallest worthwhile effect was a 30% (IQR 10 to 40) reduction in pain, additional to no intervention (Figure 3). People with a higher baseline pain intensity (0 to 10,  $\beta$  = -2.6%, 95% CI -4.2 to -1.2), longer duration of pain (weeks,  $\beta$  = -1.3%, 95% CI -2.4 to -0.2) and lower baseline level of exercise ( $\beta$  = 3.9%, 95% CI 0.6 to 7.2) required a slightly lower smallest worthwhile effect of NSAIDs. No clear associations were observed between the other baseline characteristics (ie, previous use of NSAIDs, sex and perceived health status) and the smallest worthwhile effect of NSAIDs in acute LBP (Appendix 6 on the eAddenda).

### Smallest worthwhile effect of NSAIDs for chronic LBP

Seventy (30%) participants indicated that they would not consider taking NSAIDs for their chronic LBP. For people with chronic LBP who



**Figure 2.** Flow diagram of participants in the survey.

<sup>a</sup> Participants could be ineligible for multiple reasons.

<sup>b</sup> A description of participants who did not complete the survey can be seen in Appendix 4.

would consider NSAIDs (n = 161), the smallest worthwhile effect was a 27.5% (IQR 10 to 50) pain reduction, additional to no intervention (Figure 4). There were no clear associations between any baseline characteristics and the smallest worthwhile effect of NSAIDs for people with chronic LBP (Appendix 6).

**Smallest worthwhile effect of an individualised exercise program for chronic LBP**

Twenty participants (9%) indicated that they would not consider an individualised exercise program. For people with chronic LBP who would consider exercise (n = 200), the smallest worthwhile effect was a 20% (IQR 10 to 40) reduction in pain, additional to no intervention (Figure 5). There were no clear associations between any baseline

**Table 1**  
Characteristics of participants who completed the survey.

Characteristic	Acute (n = 116)	Chronic (n = 230)
Sex, n (%) female	62 (53)	142 (62)
Pain intensity (0 to 10), mean (SD)	4.9 (2.2)	4.9 (2.3)
Age (y), mean (SD)	46.2 (14.8)	53.7 (14.5)
Duration of LBP, median (IQR)	4 wks (1.6 to 8)	13 yrs (7 to 23)
Perceived health status, n (%)		
quite poor	3 (3)	10 (4)
poor	7 (6)	22 (10)
neither good nor poor	12 (11)	52 (23)
good	51 (45)	98 (43)
quite good	41 (36)	47 (21)
Previous experience with exercise for LBP, n (%)	92 (80)	196 (85)
Previous experience with NSAIDs for LBP, n (%)	81 (70)	177 (78)
Number of health professionals seen (n), n (%)		
0	44 (38)	82 (36)
1	46 (40)	64 (28)
2	17 (15)	39 (17)
3	5 (4)	22 (10)
> 3	4 (4)	23 (10)
Not currently exercising, n (%)	18 (15)	67 (29)
Weekly physical activity (mins), n (%)		
< 30	3/98 (3)	7/163 (4)
30 to 60	17/98 (17)	19/163 (12)
60 to 90	21/98 (21)	36/163 (22)
90 to 150	19/98 (19)	37/163 (23)
> 150	36/98 (37)	62/163 (39)

LBP = low back pain, NSAID = nonsteroidal anti-inflammatory drugs.

characteristics and the smallest worthwhile effect of exercise for people with chronic LBP (Appendix 6).

Table 2 presents the hypothetical point-estimates that correspond with the smallest worthwhile effect at different levels of baseline pain.

**Sensitivity analyses**

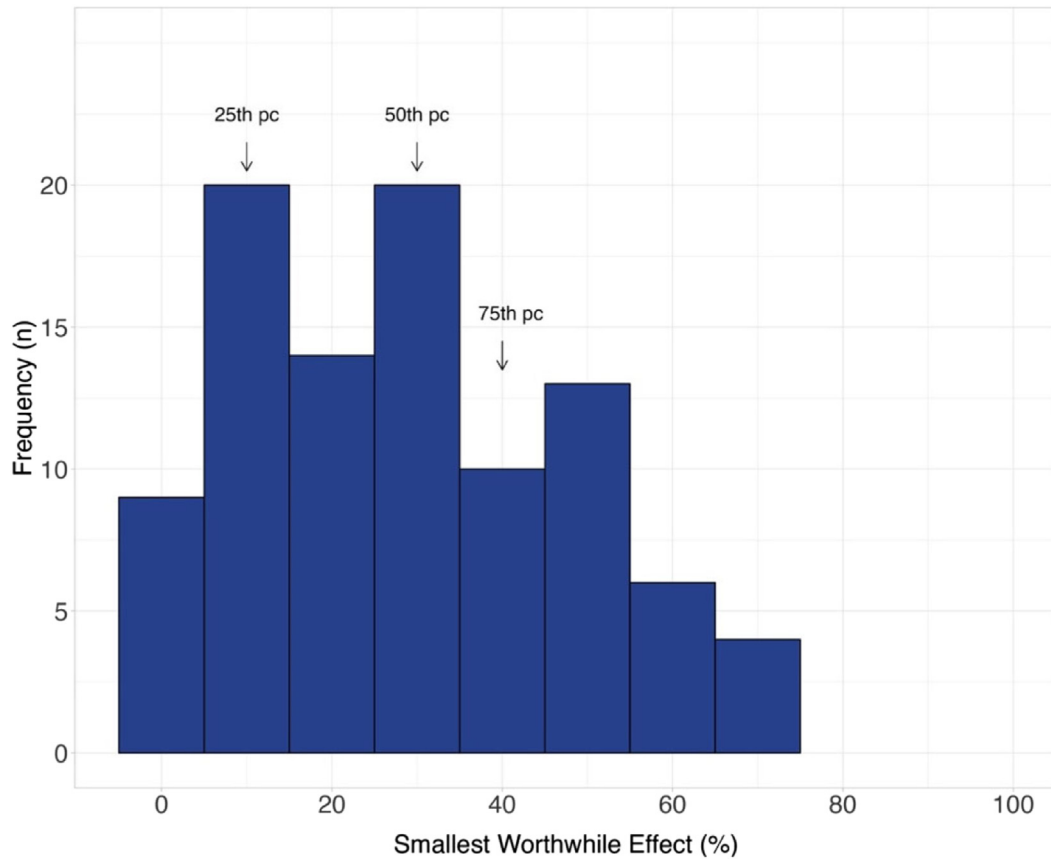
In the second survey, 689 participants accessed the survey, with 184 unique respondents providing data for at least one analysis (Appendix 7 on the eAddenda). Compared to participants with chronic LBP in the first survey, those in the second survey had: higher pain intensity (MD 0.9), shorter duration of LBP (MD -2.5 years), and poorer health status (Appendix 7).

**Smallest worthwhile effect of NSAIDs for chronic LBP**

In the second survey, where participants were shown a ‘safe’ description of NSAIDs, 41 (22%) participants indicated that they would not consider taking NSAIDs for their chronic LBP. In the ‘safe’ description of NSAIDs, for people who would consider NSAIDs (n = 142), the smallest worthwhile effect was a 30% (IQR 10 to 50%) reduction in pain, additional to no intervention (Appendix 8 on the eAddenda). The distributions were not symmetrical so we describe qualitative differences in medians. The smallest worthwhile effect was not clearly different between the two groups shown the ‘risky’ or ‘safe’ descriptions (difference in medians = 2.5%). There was a meaningful positive association between prior NSAID use and smallest worthwhile effect ( $\beta = 9.4\%$ , 95% CI 1.6 to 17.4) (ie, those who had used NSAIDs previously for their back pain required a nearly 10% greater effect of NSAIDs to consider them worthwhile).

**Smallest worthwhile effect of an individualised exercise program**

In the second survey, where participants were shown a ‘risky’ description of exercise, seven (4%) participants indicated that they would not consider an individualised exercise program. In the ‘risky’ description of exercise, for participants who would consider exercise (n = 154), the smallest worthwhile effect was a 20% (IQR 10 to 40%) reduction in pain, additional to no intervention (Appendix 9 on the eAddenda). The distributions were not symmetrical so we describe qualitative differences in medians. The smallest worthwhile effect was not different between the two groups shown the ‘risky’ or ‘safe’ description (difference in medians = 0%). Preferences for exercise modality, duration and level of supervision are displayed in Appendix 10 on the eAddenda.

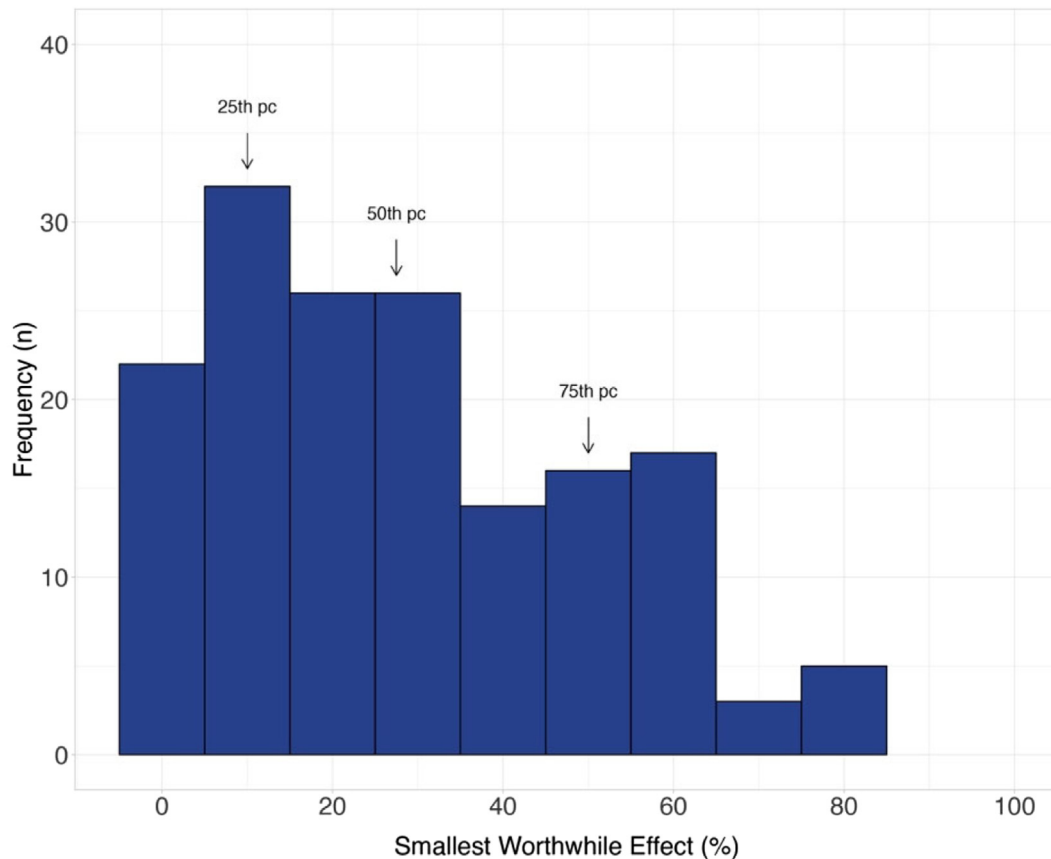


**Figure 3.** Distribution of smallest worthwhile effects of nonsteroidal anti-inflammatory drugs for people with acute low back pain (n = 96). pc = percentile.

## Discussion

This study showed that people with LBP require a 20 to 30% reduction in their pain intensity additional to no intervention (ie,

natural history, regression to the mean) to consider the effects of common guideline-recommended interventions (ie, NSAIDs and exercise) worthwhile. Irrespective of pain duration (acute or chronic), people with LBP require an approximately 30% reduction in pain intensity for



**Figure 4.** Distribution of smallest worthwhile effects of nonsteroidal anti-inflammatory drugs for people with chronic low back pain (n = 161). pc = percentile.

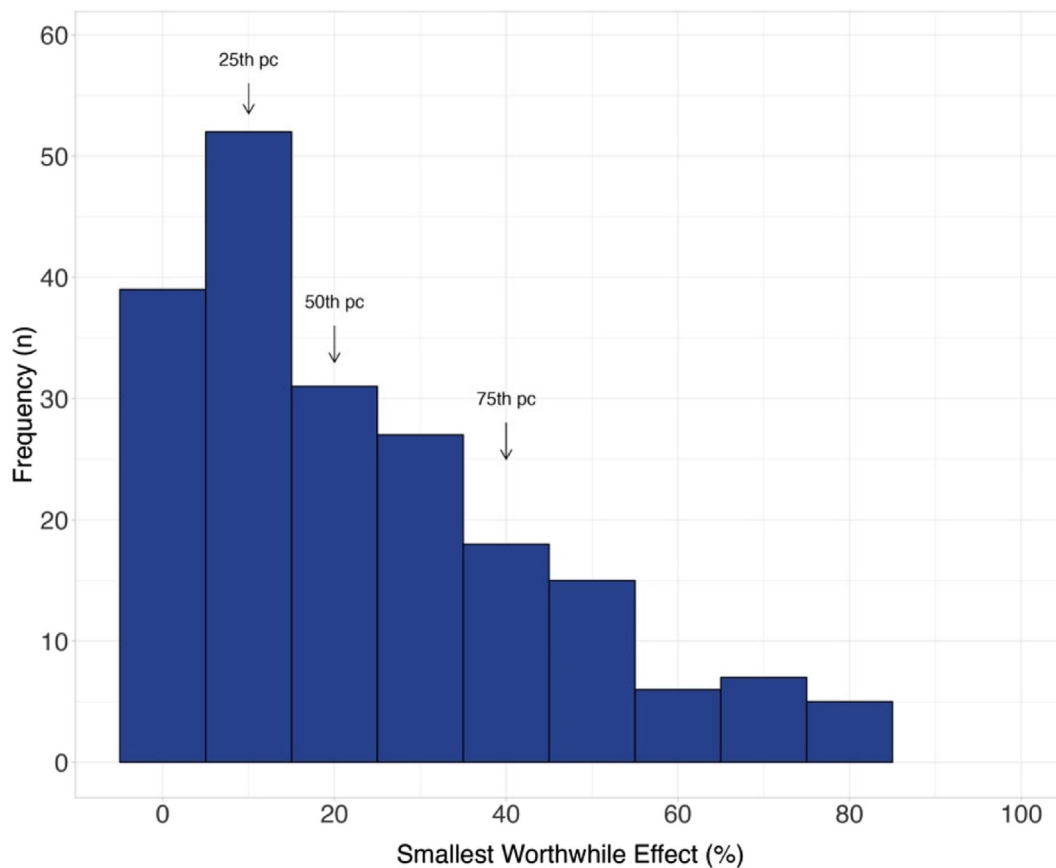


Figure 5. Distribution of smallest worthwhile effects for exercise for people with chronic low back pain (n = 200). pc = percentile.

NSAIDs to be worthwhile. The smallest worthwhile effect for exercise is lower at a 20% reduction in pain. For people with acute LBP, the type of participants who had a higher pain intensity, longer duration of back pain and lower level of current exercise required a lower smallest worthwhile effect. No baseline characteristics were clearly associated with requiring a lower or higher smallest worthwhile effect for people with chronic LBP. The mean pain intensity for this sample was 4.9 out of 10 for acute and chronic LBP, which was similar to or slightly lower than samples of care-seeking adults with LBP in cohort studies<sup>31</sup> and several large randomised trials and meta-analyses.<sup>11,48–52</sup> We believe that our results are generalisable to those seeking care.

Efforts to estimate the smallest worthwhile effect are rarely reported in the literature.<sup>15</sup> For LBP, the smallest worthwhile effect of NSAIDs and physiotherapy have previously been estimated<sup>14,19</sup> at a 30% and 20% reduction in pain, respectively. We found a smallest worthwhile effect of 27.5% of NSAIDs for people with chronic LBP, similar to

that of Ferreira et al,<sup>14</sup> increasing confidence that the true value lies at approximately 30%. We also extended these findings to people with acute LBP, showing similar results (30% reduction in pain). Despite similar smallest worthwhile effects, more participants with chronic LBP (n = 70, 30%) would not consider taking NSAIDs, compared with those with acute LBP (n = 20, 18%). This finding may be because people with chronic low back pain had more experiences with NSAIDs (chronic n = 177, 78%; acute n = 81, 70%) where NSAIDs had not helped to relieve their pain. For exercise, the smallest worthwhile effect was the same as reported for physiotherapy (20% reduction in pain), indicating that these non-pharmacological interventions may be considered similarly worthwhile. Our findings enhance confidence in the existing estimates of the smallest worthwhile effect but also extend these to include estimates for currently recommended interventions.

All estimates of the smallest worthwhile effect are specific to the comparison studied, in our case, NSAIDs or exercise compared with no intervention. Our study demonstrates the importance of intervention-specific measures when determining whether intervention effects are meaningful to patients. In our study, and previously, the smallest worthwhile effect for pharmacological interventions is higher than that for non-pharmacological ones.<sup>14</sup> However, we aimed to mitigate the potential generalisability concerns of previous studies that recruited via physiotherapy clinics by sampling the general community via social media. Despite this, most of our sample reported that they currently exercised, which may have resulted in a group of people who prefer exercise as a modality to manage their LBP. It may also be that participants' preconceptions about the safety and other benefits of exercise such as developing a healthy lifestyle<sup>53</sup> may be a reason for the lower smallest worthwhile effect for exercise, compared with NSAIDs, even when the risks of NSAIDs were described as being minor. Further research into the determinants of the smallest worthwhile effect is required.

Despite being a widely used method to estimate the smallest worthwhile effect,<sup>14,19</sup> the measurement properties of the benefit-harm trade-off method are not well established. Nonetheless, benefit-harm

Table 2

Point estimates of the smallest worthwhile effect corresponding with different hypothetical levels of baseline pain on the 0-to-10 numerical pain rating scale.

Hypothetical baseline pain (0 to 10)	Point estimate of the SWE in addition to no intervention (0 to 10) <sup>a</sup>		
	Acute LBP (NSAIDs <sup>b</sup> )	Chronic LBP (NSAIDs <sup>c</sup> )	Chronic LBP (Exercise <sup>d</sup> )
4	1.2	1.1	0.8
5 <sup>e</sup>	1.5	1.4	1
6	1.8	1.7	1.2
7	2.1	1.9	1.4
8	2.4	2.2	1.6

LBP = low back pain, NSAIDs = nonsteroidal anti-inflammatory drugs, SWE = smallest worthwhile effect.

<sup>a</sup> The point estimate is the reduction in pain required, in addition to no intervention.

<sup>b</sup> The smallest worthwhile effect of NSAIDs for acute LBP was 30%.

<sup>c</sup> The smallest worthwhile effect of NSAIDs for chronic LBP was 27.5%.

<sup>d</sup> The smallest worthwhile effect for exercise for chronic LBP was 20%.

<sup>e</sup> Baseline pain in our sample.

trade-off has notable strengths compared with other methods (eg, anchor-based or distribution-based methods)<sup>54</sup> often used to determine the clinical importance of an intervention's effect. The benefit-harm trade-off method, unlike other methods, forces participants to choose between two hypothetical scenarios, where they receive a treatment and where they do not; the smallest worthwhile effect must therefore be interpreted as a between-group difference. Benefit-harm trade-off also relies on the patients' interpretation of the intervention(s) described; they will inevitably be influenced by prior beliefs and experience. To understand the determinants of the smallest worthwhile effect, we conducted an exploratory sensitivity analysis that described NSAIDs in safer terms and exercise in riskier terms than in the original study. Our findings suggest that the description of harms does not alter the smallest worthwhile effect. The discrete choice experiment is another method used to estimate the smallest worthwhile effect, particularly for complex interventions such as exercise, where multiple characteristics may change how worthwhile an intervention may be. The benefit-harm trade-off method is limited in this regard, as only the amount of benefit changes across the decision-making process.<sup>13</sup> However, estimates obtained from discrete choice experiments have shown discrepancies with estimates from benefit-harm trade-off studies and thus the validity of the discrete choice experiment is uncertain.<sup>20</sup> For the smallest worthwhile effect to be adopted more broadly, the field may benefit from assessing the reliability and validity of the benefit-harm trade-off and discrete choice experiment methods.

To help patients easily rate improvements in their pain intensity in the benefit-harm trade-off questions, we presented a percentage reduction in pain intensity as opposed to an absolute change (eg, 2 out of 10). Using a percentage has the benefit of the smallest worthwhile effect being easily applied to any quantitative scales of a similar construct (as done in Table 2), not just the one investigated. However, to further strengthen the validity of the benefit-harm trade-off method, qualitative work should be conducted to identify how best to ask patients these questions. Future research may benefit from investigating the smallest worthwhile effect of exercise and NSAIDs on scales of other constructs such as physical function. We followed a similar approach to previous research and found that our estimates of the smallest worthwhile effect were similarly distributed,<sup>14</sup> with an interquartile range of 10 to 40 for NSAIDs and range spanning all options (0 to 70% for acute LBP, 0 to 80% for chronic LBP). The range of possible values does not extend to 100% as the smallest worthwhile effect represents a between-group difference and must not be interpreted as a change from baseline.

Our study extends previous research by investigating currently recommended interventions for acute and chronic LBP, providing researchers, clinicians and policymakers with up-to-date evidence to guide the interpretation of results from randomised trials and meta-analyses of NSAIDs and exercise when compared with no intervention. When interpreting a meta-analysis of exercise compared with no treatment, usual care or placebo with the smallest worthwhile effect, exercise clearly produces clinically meaningful reductions in pain intensity for chronic LBP (-1.52 out of 10 (95% CI -1.83 to -1.22), equating to a difference of -30% between groups).<sup>11</sup> Exercise is seldom compared with placebo/sham in clinical trials of LBP<sup>11</sup> so our results for the smallest worthwhile effect for exercise may have greater applicability for interpreting these studies.

Our study used the benefit-harm trade-off method<sup>16</sup> to establish the smallest worthwhile effect.<sup>12</sup> In contrast to widely used clinical significance measures like the minimum clinically important difference,<sup>15</sup> the smallest worthwhile effect is determined by patients, is specific to an intervention accounting for its costs, risks and inconveniences, and represents a between-group difference. These characteristics of the smallest worthwhile effect enable a more patient-centred interpretation of randomised trials and meta-analyses<sup>12,13,22</sup> of interventions for which the smallest worthwhile effect has been determined.

Our sample was recruited via social media, which may have limited generalisability.<sup>55,56</sup> However, our sample does appear similar to care-seeking LBP populations in terms of pain intensity.<sup>11,31,48-50</sup> To

determine the smallest worthwhile effect, we have provided a comparison between intervention and no intervention. This is not always reflected in clinical trials of pharmacological interventions, where a placebo comparison is more common. The 'effect' of placebo interventions may be greater than no intervention for people with LBP,<sup>57</sup> making it more challenging to apply the results of this study to the interpretation of between-group effects in placebo-controlled trials. In our post-hoc comparison between the different descriptions of the harms of NSAIDs and exercise for chronic LBP, participants were not randomised to either the 'safe' or 'risky' descriptions of the interventions; therefore, differences between the groups of participants may have masked the true effect of the description of harms of each intervention. Because of this limitation, we cannot completely exclude the possibility that described harms impact the smallest worthwhile effect.

Few estimates of the smallest worthwhile effect have been published across medical and health research,<sup>17,18,20,21</sup> including estimates for only two interventions for chronic LBP<sup>14,19</sup> in the field of pain, all compared with no intervention. There is the opportunity to estimate the smallest worthwhile effect of recommended interventions compared with other commonly used 'active' interventions (ie, opioids, anti-depressants, muscle relaxants or surgeries). These estimates are needed to interpret the comparative effectiveness of these interventions and guide policy recommendations. Further, the field may benefit from consensus on the approach to sample size estimation, as currently there are no guidelines for determining the sample required in benefit-harm trade-off.

Face validity can be assumed for the benefit-harm trade-off directly, as it measures the smallest effect that a patient deems to be worthwhile. However, it is unclear which design variables of the benefit-harm trade-off may affect the content validity and reliability of the smallest worthwhile effect, warranting further investigation. With evidence of sufficient reliability and validity, researchers who test interventions should consider including a benefit-harm trade-off study within their trials to improve interpretation of their findings.

The smallest worthwhile effect of NSAIDs is a 30% reduction in pain for acute LBP and a 27.5% reduction in pain for chronic LBP, additional to no intervention. The smallest worthwhile effect for an individualised exercise program for chronic LBP is a 20% reduction in pain, additional to no intervention. These estimates are determined by patients and can be used to aid interpretation of randomised trials and meta-analyses. The field would benefit from further research exploring the validity of the benefit-harm trade-off method to estimate the smallest worthwhile effect as well as the smallest worthwhile effect for different comparisons (eg, comparisons with other commonly used pharmacological or surgical interventions for low back pain).

**What was already known on this topic:** The smallest worthwhile effect, determined with benefit-harm trade-off, represents the minimum between-group difference in an outcome that a patient considers worth the costs, risks and inconveniences of an intervention, in addition to a comparator. The smallest worthwhile effect is the most appropriate index of clinical importance to inform treatment decision-making and interpret the findings of clinical trials and meta-analyses. The smallest worthwhile effect for pain intensity has been estimated for 'physiotherapy' and nonsteroidal anti-inflammatory drugs for people with chronic low back pain as being 20% and 30%, respectively, compared with no treatment.

**What this study adds:** The smallest worthwhile effect for NSAIDs was a 30% and 27.5% reduction in pain, additional to no treatment, for people with acute and chronic low back pain, respectively. The smallest worthwhile effect for an individualised exercise program for people with chronic low back pain was a 20% reduction in pain, additional to no treatment. It does not appear that changing the way risks of these treatments are described to people with chronic low back pain affects the smallest worthwhile effect.

**Footnotes:** <sup>a</sup> R version 4.2.0, RCoreTeam, Vienna, Austria.

**eAddenda:** Appendices 1 to 10 can be found online at <https://doi.org/10.1016/j.jphys.2023.08.006>

**Ethics approval:** The study was approved by the University of New South Wales Human Research Ethics Committee (HC200982).

**Competing interests:** Nil.

**Source(s) of support:** There was no specific funding for this study. HJH was supported by an Australian National Health and Medical Research Council (NHMRC) Postgraduate Scholarship, a PhD Top-Up Scholarship from Neuroscience Research Australia, and was a Neuroscience Research Australia PhD Pearl sponsored by Sandra Salteri. AGC was supported by an Australian NHMRC Investigator Grant (ID 2010088). JMH was supported by an Australian NHMRC Investigator Grant (ID 2010128). SS is supported by the International Association for the Study of Pain (IASP) John J. Bonica Postdoctoral Fellowship.

**Acknowledgements:** We wish to acknowledge all participants for their valued contribution to this work.

**Provenance:** Not invited. Peer reviewed.

**Data availability:** All data and code will be shared publicly on OSF upon publication ([osf.io/3erjx/](https://osf.io/3erjx/)).

**Correspondence:** Harrison J Hansford, Centre for Pain IMPACT, Neuroscience Research Australia, Sydney, Australia. Email: [h.hansford@unsw.edu.au](mailto:h.hansford@unsw.edu.au)

## References

- Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990, 2013, 2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*. 2020;396:1204–1222. [https://doi.org/10.1016/S0140-6736\(20\)30925-9](https://doi.org/10.1016/S0140-6736(20)30925-9)
- Walker BF, Muller R, Grant WD. Low back pain in Australian adults: health provider utilization and care seeking. *J Manipulative Physiol Ther*. 2004;27:327–335. <https://doi.org/10.1016/j.jmpt.2004.04.006>
- Chiarotto A, Koes BW. Nonspecific low back pain. *N Engl J Med*. 2022;386:1732–1740. <https://doi.org/10.1056/NEJMc2032396>
- Oliveira CB, Maher CG, Pinto RZ, Traeger AC, Lin CW, Chenot JF, et al. Clinical practice guidelines for the management of non-specific low back pain in primary care: an updated overview. *Eur Spine J*. 2018;27:2791–2803. <https://doi.org/10.1007/s00586-018-5673-2>
- Ahern M, Dean CM, Dear BF, Willcock SM, Hush JM. The experiences and needs of people seeking primary care for low-back pain in Australia. *PAIN Reports*. 2019;4:e756. <https://doi.org/10.1097/pr9.0000000000000756>
- Hush JM, Refshauge K, Sullivan G, De Souza L, Maher CG, McAuley JH. Recovery: what does this mean to patients with low back pain? *Arthritis Rheum*. 2009;61:124–131. <https://doi.org/10.1002/art.24162>
- Chiarotto A, Deyo RA, Terwee CB, Boers M, Buchbinder R, Corbin TP, et al. Core outcome domains for clinical trials in non-specific low back pain. *Eur Spine J*. 2015;24:1127–1142. <https://doi.org/10.1007/s00586-015-3892-3>
- McRae M, Hancock MJ. Adults attending private physiotherapy practices seek diagnosis, pain relief, improved function, education and prevention: a survey. *J Physiother*. 2017;63:250–256. <https://doi.org/10.1016/j.jphys.2017.08.002>
- van der Gaag WH, Roelofs PDDM, Enthoven WTM, van Tulder MW, Koes BW. Non-steroidal anti-inflammatory drugs for acute low back pain. *Cochrane Database Syst Rev*. 2020;4:CD013581. <https://doi.org/10.1002/14651858.CD013581>
- Enthoven WTM, Roelofs P, Deyo RA, van Tulder MW, Koes BW. Non-steroidal anti-inflammatory drugs for chronic low back pain. *Cochrane Database Syst Rev*. 2016;2:CD012087. <https://doi.org/10.1002/14651858.CD012087>
- Hayden JA, Ellis J, Ogilvie R, Malmivaara A, van Tulder MW. Exercise therapy for chronic low back pain. *Cochrane Database Syst Rev*. 2021;9:CD009790. <https://doi.org/10.1002/14651858.CD009790.pub2>
- Ferreira ML, Herbert RD. What does 'clinically important' really mean? *Aust J Physiother*. 2008;54:229–230. [https://doi.org/10.1016/s0004-9514\(08\)70001-1](https://doi.org/10.1016/s0004-9514(08)70001-1)
- Ferreira M. Research Note: The smallest worthwhile effect of a health intervention. *J Physiother*. 2018;64:272–274. <https://doi.org/10.1016/j.jphys.2018.07.008>
- Ferreira ML, Herbert RD, Ferreira PH, Latimer J, Ostelo RW, Grotle M, et al. The smallest worthwhile effect of nonsteroidal anti-inflammatory drugs and physiotherapy for chronic low back pain: a benefit-harm trade-off study. *J Clin Epidemiol*. 2013;66:1397–1404. <https://doi.org/10.1016/j.jclinepi.2013.02.018>
- Ferreira ML, Herbert RD, Ferreira PH, Latimer J, Ostelo RW, Nascimento DP, et al. A critical review of methods used to determine the smallest worthwhile effect of interventions for low back pain. *J Clin Epidemiol*. 2012;65:253–261. <https://doi.org/10.1016/j.jclinepi.2011.06.018>
- Barrett B, Brown D, Mundt M, Brown R. Sufficiently important difference: expanding the framework of clinical significance. *Med Decis Making*. 2005;25:250–261. <https://doi.org/10.1177/0272989x05276863>
- Barrett B, Brown R, Mundt M, Dye L, Alt J, Safdar N, et al. Using benefit harm tradeoffs to estimate sufficiently important difference: the case of the common cold. *Med Decis Making*. 2005;25:47–55. <https://doi.org/10.1177/0272989x04273147>
- Barrett B, Harahan B, Brown D, Zhang Z, Brown R. Sufficiently important difference for common cold: severity reduction. *Ann Fam Med*. 2007;5:216–223.
- Christiansen DH, de Vos Andersen NB, Poulsen PH, Ostelo RW. The smallest worthwhile effect of primary care physiotherapy did not differ across musculoskeletal pain sites. *J Clin Epidemiol*. 2018;101:44–52. <https://doi.org/10.1016/j.jclinepi.2018.05.019>
- Franco MR, Howard K, Sherrington C, Rose J, Ferreira PH, Ferreira ML. Smallest worthwhile effect of exercise programs to prevent falls among older people: estimates from benefit-harm trade-off and discrete choice methods. *Age Ageing*. 2016;45:806–812. <https://doi.org/10.1093/ageing/afw110>
- McNamara RJ, Elkins MR, Ferreira ML, Spencer LM, Herbert RD. Smallest worthwhile effect of land-based and water-based pulmonary rehabilitation for COPD. *ERJ Open Res*. 2015;1:00007–2015. <https://doi.org/10.1183/23120541.00007-2015>
- Kamper SJ. Interpreting Outcomes 3—Clinical Meaningfulness: Linking Evidence to Practice. *J Orthop Sports Phys Ther*. 2019;49:677–678. <https://doi.org/10.2519/jospt.2019.0705>
- Copay AG, Subach BR, Glassman SD, Polly DW, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J*. 2007;7:541–546. <https://doi.org/10.1016/j.spinee.2007.01.008>
- Gatchel RJ, Mayer TG. Testing minimal clinically important difference: consensus or conundrum? *Spine J*. 2010;10:321–327. <https://doi.org/10.1016/j.spinee.2009.10.015>
- Foster ED, Deardorff A. Open science framework (OSF). *J Med Libr Assoc*. 2017;105:203.
- Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Bull World Health Org*. 2007;85:867–872.
- Ware Jr JE. SF-36 health survey update. *Spine*. 2000;25:3130–3139.
- Chiarotto A, Maxwell LJ, Ostelo RW, Boers M, Tugwell P, Terwee CB. Measurement Properties of Visual Analogue Scale, Numeric Rating Scale, and Pain Severity Subscale of the Brief Pain Inventory in Patients With Low Back Pain: A Systematic Review. *J Pain*. 2019;20:245–263. <https://doi.org/10.1016/j.jpain.2018.07.009>
- Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform*. 2019;95:103208. <https://doi.org/10.1016/j.jbi.2019.103208>
- Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42:377–381. <https://doi.org/10.1016/j.jbi.2008.08.010>
- Costa LD, Maher CG, Hancock MJ, McAuley JH, Herbert RD, Costa LOP. The prognosis of acute and persistent low-back pain: a meta-analysis. *Can Med Assoc J*. 2012;184:E613–E624. <https://doi.org/10.1503/cmaj.11271>
- Hush JM, Refshauge K, Sullivan G, De Souza L, Maher CG, McAuley JH. Recovery: What does this mean to patients with low back pain? *Arthritis Care Res*. 2009;61:124–131. <https://doi.org/10.1002/art.24162>
- Froud R, Patel S, Rajendran D, Bright P, Bjørkli T, Buchbinder R, et al. A systematic review of outcome measures use, analytical approaches, reporting methods, and publication volume by year in low back pain trials published between 1980 and 2012: Respite, adipscie, et prospice. *PLoS One*. 2016;11:e0164573. <https://doi.org/10.1371/journal.pone.0164573>
- Kelley K, Maxwell SE. Sample size for multiple regression: obtaining regression coefficients that are accurate, not simply significant. *Psychol Meth*. 2003;8:305.
- R: A language and environment for statistical computing. R Foundation for Statistical Computing; 2020. <https://www.R-project.org/>
- Yoshida K, Bohn J, Yoshida MK. Package 'tableone'. *R Foundation for Statistical Computing, Vienna, Austria (30 November 2016)*. 2020.
- Wickham H. *The tidyverse. R package ver. 2017;1:836*.
- Walker A, Braglia L. Package 'openxlsx'. 2018.
- Canty AJ. Resampling methods in R: the boot package. *The Newsletter of the R Project Volume*. 2002;2:2–7.
- Read: *Read rectangular text data*. 2018.
- Wickham H, Chang W, Wickham MH. Package 'ggplot2. *Create elegant data visualisations using the grammar of graphics Version*. 2016;2:1–189.
- Kassambara A, Kassambara MA. Package 'ggpubr'. *R package version 01*. 2020;6.
- R Core Team. R: A language and environment for statistical computing. 2013;
- Efron B, Tibshirani RJ. *An introduction to the bootstrap*. Florida, USA: CRC press; 1994.
- Hancock MJ, Kent P. Letter re: Christiansen DH, de Vos Andersen NB, Poulsen PH, Ostelo RW. The smallest worthwhile effect of primary care physiotherapy did not differ across musculoskeletal pain sites. *J Clin Epidemiol*. 2019;105:147.
- Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). *Biometrika*. 1965;52:591–611.
- Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. *Ann Math Stat*. 1947;50–60.
- Bagg MK, Wand BM, Cashin AG, Lee H, Hübscher M, Stanton TR, et al. Effect of graded sensorimotor retraining on pain intensity in patients with chronic low back pain: a randomized clinical trial. *JAMA*. 2022;328:430–439. <https://doi.org/10.1001/jama.2022.9930>
- Mathieson S, Maher CG, McLachlan AJ, Latimer J, Koes BW, Hancock MJ, et al. Trial of Pregabalin for Acute and Chronic Sciatica. *New Eng J Med*. 2017;376:1111–1120. <https://doi.org/10.1056/NEJMoa1614292>
- Williams CM, Maher CG, Latimer J, McLachlan AJ, Hancock MJ, Day RO, et al. Efficacy of paracetamol for acute low-back pain: a double-blind, randomised controlled trial. *Lancet*. 2014;384:1586–1596. [https://doi.org/10.1016/S0140-6736\(14\)60805-9](https://doi.org/10.1016/S0140-6736(14)60805-9)
- Hancock MJ, Maher CG, Latimer J, McLachlan AJ, Cooper CW, Day RO, et al. Assessment of diclofenac or spinal manipulative therapy, or both, in addition to recommended first-line treatment for acute low back pain: a randomised controlled trial. *Lancet*. 2007;370:1638–1643. [https://doi.org/10.1016/s0140-6736\(07\)61686-9](https://doi.org/10.1016/s0140-6736(07)61686-9)



52. Kent P, Haines T, O'Sullivan P, Smith A, Campbell A, Schutze R, et al. Cognitive functional therapy with or without movement sensor biofeedback versus usual care for chronic, disabling low back pain (RESTORE): a randomised, controlled, three-arm, parallel group, phase 3, clinical trial. *Lancet*. 2023;401:1866–1877. [https://doi.org/10.1016/S0140-6736\(23\)00441-5](https://doi.org/10.1016/S0140-6736(23)00441-5)
53. Becker WC, Dorflinger L, Edmond SN, Islam L, Heapy AA, Fraenkel L. Barriers and facilitators to use of non-pharmacological treatments in chronic pain. *BMC Fam Pract*. 2017;18:41. <https://doi.org/10.1186/s12875-017-0608-2>
54. Wright A, Hannon J, Hegedus EJ, Kavchak AE. Clinimetrics corner: a closer look at the minimal clinically important difference (MCID). *J Man Manip Ther*. 2012;20:160–166. <https://doi.org/10.1179/2042618612y.0000000001>
55. Benedict C, Hahn AL, Diefenbach MA, Ford JS. Recruitment via social media: advantages and potential biases. *Digit Health*. 2019;5:2055207619867223. <https://doi.org/10.1177/2055207619867223>
56. Arigo D, Pagoto S, Carter-Harris L, Lillie SE, Nebeker C. Using social media for health research: Methodological and ethical considerations for recruitment and intervention delivery. *Digit Health*. 2018;4:2055207618771757. <https://doi.org/10.1177/2055207618771757>
57. Strijkers RHW, Schreijenberg M, Gerger H, Koes BW, Chiarotto A. Effectiveness of placebo interventions for patients with nonspecific low back pain: a systematic review and meta-analysis. *Pain*. 2021;162:2792–2804. <https://doi.org/10.1097/j.pain.0000000000002272>