ORIGINAL ARTICLE

The smallest worthwhile change on function from a self-management intervention for non-persistent low back pain

Carolina Gassen Fritsch[1] · Paulo H. Ferreira[2] · Thomas Lung[3] · Andrew J. McLachlan[4] · Manuela L. Ferreira[1]

Abstract

Purpose To determine: (1) the smallest change in function patients would need to see following a self-management intervention for low back pain (LBP) to consider it worthwhile; (2) the association between patient-related factors and the magnitude of the smallest worthwhile change.

Methods A cross-sectional analysis of 212 participants of the TEXT4myBACK randomised trial was conducted. At baseline, participants nominated the smallest change in function (0–30 scale) following a self-management program they would need to reach to consider it worthwhile. A multivariate regression model estimated the effects of demographic, comorbidities, lifestyle and LBP-related factors on the smallest worthwhile change estimates.

Results On average, people with LBP need to experience an improvement of at least 9.4 points (SD: 5.7) in function to consider a self-management intervention worthwhile. Only baseline function severity was significantly associated with the smallest worthwhile estimate (-0.60; 95%CI -0.76, -0.44).

Conclusion On average, an improvement of 9.4 points (or 31%) in function is considered by people with LBP as the smallest change that makes self-management worthwhile. Those with lower levels of function needed to experience greater improvements.

Keywords Sufficiently important difference · Low back pain · Self-management · Function

✉ Carolina Gassen Fritsch
carolina.gassenfritsch@sydney.edu.au

Paulo H. Ferreira
paulo.ferreira@sydney.edu.au

Thomas Lung
tom.lung@sydney.edu.au

Andrew J. McLachlan
andrew.mclachlan@sydney.edu.au

Manuela L. Ferreira
manuela.ferreira@sydney.edu.au

1 Sydney Musculoskeletal Health, The Kolling Institute, School of Health Sciences, Faculty of Medicine and Health, The University of Sydney, Level 10, Kolling Building, Royal North Shore Hospital, Reserve Road, St Leonards, Sydney, NSW 2065, Australia

2 Musculoskeletal Health Research Group, Sydney Musculoskeletal Health, School of Health Sciences, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia

3 Sydney School of Public Health, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia

4 Sydney Pharmacy School, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia

Introduction

Randomised controlled trials are conducted to establish the effectiveness of different interventions on various health outcomes [1]. Analysis of the effects of interventions may include null hypothesis testing and estimation of the size of the effect [1]. However, the effect of an intervention on an outcome of interest may be statistically significant when compared to a control intervention, but fail to reach clinical relevance or significance [1].

Arguably, the definition of clinical significance should be based on judgements of healthcare consumers and should be specific to the intervention of interest [2]. It should also be elicited in a way that allows its users to appraise treatment effects, i.e. the differences in outcomes between the intervention of interest and the control intervention [3]. Since the smallest worthwhile effect is specific to a population and an intervention, it would be arguably randomised controlled trial-specific too. Randomised controlled trials investigate the effects of one intervention (with varied characteristics) compared to a control intervention on a population

of specific clinical and demographic characteristics, and therefore require a specific smallest worthwhile effect. In the absence of trial-specific estimates, most researchers will refer to existing estimates, a commonly used one being the minimal clinically important difference or minimal clinically important improvement. These estimates are elicited by anchor-based approaches by associating a change in the outcome with some other subjective assessment of improvement, such as the global rating scale [2, 4]. The threshold for eliciting the minimal clinically important change, however, is chosen by researchers and not patients [2, 5]. Therefore, these methods have been criticised for omitting the perspective of patients or consumers and failing to account for the specific risks, costs and inconveniences of an intervention in the estimating process [2].

Other methods have been used in an attempt to estimate thresholds of clinical significance for the effects of interventions, such as the benefit-harm trade-off approach [5] and discrete choice experiments [6]. These methods have been recommended as gold-standard methodologies to estimate the threshold of clinical significance of treatment effects as they allow researchers to elicit—based on consumers' perspectives—the smallest difference in an outcome between an intervention and a control, that would make that intervention worth its risks, costs and inconveniences [3]. However, there are some barriers to using these approaches, including time and resource commitments.

In the attempt of overcoming past limitations, we have employed a modified benefit-harm trade-off approach that is simpler and less burdensome to the participant (than the methodology used to elicit the smallest worthwhile effect) and could be incorporated into the data collection process of randomised trials. A short question added to the baseline survey of a randomised clinical trial can be used to elicit the smallest worthwhile change from the participants' perspective. This question would explain the possible risks, harms and inconveniences expected from the intervention and ask participants which would be the smallest change or improvement in a health outcome they would need to reach at the end of the intervention to consider it worthwhile. It would allow a fast way to elicit the smallest worthwhile change that could be used in a responder analysis.

Thus, this study aimed to: (1) estimate the smallest worthwhile change needed for a self-management intervention consisting of text messages for non-persistent, non-specific low back pain to be considered worthwhile; (2) investigate if demographic characteristics, comorbidities, lifestyle factors and low back pain clinical characteristics were associated with the magnitude of the smallest worthwhile change. We have used data from the TEXT4myBACK trial [7]. TEXT-4myBACK is a randomised controlled trial investigating the effects of a self-management text message intervention compared to control on function of people with non-specific,

non-persistent low back pain [7]. The TEXT4myBACK clinical trial was approved by the Northern Sydney Local Health District Ethics Committee in Australia (ETH 13895) [7].

Material and methods

This study is a cross-sectional analysis of baseline data from 212 participants of the TEXT4myBACK randomised controlled trial [7]. Community-dwelling adults with low back pain living in Australia were invited to participate in the TEXT4myBACK clinical trial [7]. People aged 18 years or older who had an episode of non-specific low back pain for less than 12 weeks, with or without the presence of leg pain, had pain classified at least as 'moderate' on the SF-12 pain scale [8] and had familiarity with the use and access to, a telephone that can receive text messages were included. Pregnant women, people who had spinal surgery within the preceding year or symptoms that could indicate radiculopathy or a serious spinal pathology, co-morbid health conditions that prevented active participation in physical activity programs, inadequate English to understand the text messages or complete the study surveys or any disorder that reduced their ability to understand and give informed consent were excluded.

People who met the criteria and signed the online consent form were included in the TEXT4myBACK trial. Participants completed an online questionnaire in the REDCap software [9], which included questions on demographic characteristics, comorbidities, low back pain clinical profile, pain intensity, function, physical activity participation, sedentary behaviour, and eHealth literacy.

Physical function was assessed with the Patient-Specific Functional Scale [10]. Each participant was asked to name three important activities they were unable to do or had difficulties in performing due to their low back pain. They scored each activity using a numerical rating scale from 0 to 10, where 0 meant unable to perform activity and 10 meant able to perform activity at the pre-injury level. The scores were summed, and their individual total function score was presented (ranging from 0 to 30 points).

Following this question in the baseline survey, each participant was asked to nominate the smallest score on this function scale they would need to achieve to consider a self-management intervention worthwhile. A short description of self-management along with costs and inconveniences was provided (Box 1). However, some intervention attributes (such as the number of text messages being sent, their frequency and time) were kept from participants to ensure blinding once randomised.

Predictors

Demographic characteristics (i.e., age, gender and educational level), comorbidities, lifestyle factors (i.e., self-reported sleep issues and sedentary behaviour) and low back pain clinical profile (i.e., pain intensity, function, presence of leg pain, pain duration and quality of life) were prospectively chosen as predictors. Comorbidities were assessed with the Self-Administered Comorbidity Questionnaire [11]. Sleep issues were self-reported difficulty in falling asleep or waking up at night. Sedentary behaviour was assessed with the Sedentary Behaviour Questionnaire [12]. Pain intensity was assessed as the average pain intensity in the previous week on a 0–100 visual analogue scale, where 0 was no pain and 100 was the worst pain ever [13]. Quality of life was evaluated with the ED-5Q-5L questionnaire [14].

Statistical analysis

Power analysis

Sample size calculations were conducted with G*Power software (version 3.1.9.2) to ascertain study power. Based on a priori sample size calculation, a minimum sample of 178 participants would be required to assess the association of eleven predictors with the estimates of smallest worthwhile change, with power of 0.95 at an alpha error level of 0.05.

Data analysis

Baseline demographic data and the distribution of the smallest worthwhile change of participants of both groups will be presented by central tendency (mean and median) and variability (standard deviation—SD, 25th and 75th percentile or range). Missing data or drop-outs were not included in the analyses, as they would not represent individualised values.

A multiple linear regression model was used to quantify the effect of the predictors on the magnitude of the smallest worthwhile change scores, and 95% confidence intervals were calculated. The assumptions of the linear model were assessed by performing residual analysis. The modified Breusch–Pagan (BP) test for heteroscedasticity was used to assess the assumption of constant error variance. If the results of the BP test indicated non-constant error variance, robust heteroscedastic consistent standard errors were used. The normality of the error was assessed with Q–Q plots and Kolmogorov–Smirnov test based on model residuals. Standardised residuals greater than the absolute value of 2.0 were considered outliers for secondary analysis. If outliers were detected, a secondary multivariate linear model was conducted without the outliers. Partial eta squared ($\eta2$) measures the proportion of the total variance in the outcome explained by an independent variable and after accounting for the variance explained by other variables in the model. Partial eta squared was considered to interpret the magnitude of the effect of each predictor,
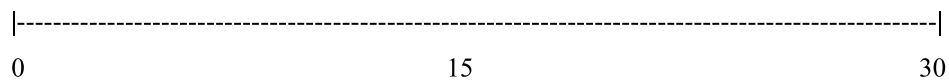
Box 1 Smallest worthwhile change question

> *Smallest worthwhile effect on function*
>
> Self-management may include recommendations to remain active, read educational material or apply superficial heat. There is usually no cost involved in self-management and there is some time commitment depending on the time needed to read the information and follow the recommendations.
>
> Based on your current score of [sum_function]/30 points, please place the slider below on the **SMALLEST** (lowest) function score you would need to reach at the end of this self-management intervention to consider it worthwhile.
>
> * Please note that the SMALLEST (lowest) score should be equal to or greater than your current score.
>
> |--|
> 0   15   30
>
> 0 – unable to perform; 30 – able to perform activities to the same level as before the back pain

where $\eta2 = 0.02$ was considered small, $\eta2 = 0.13$ as moderate and $\eta2 = 0.26$ as large effect [15]. All the statistical analyses were performed at a 0.05 level of significance ($p < 0.05$) using the SPSS software (version 28).

Results

The demographic characteristics, general health and low back pain clinical profile of the 212 participants included in the current study are presented in Table 1. Data from further five participants were available but could not be included in the analysis since they misunderstood either the function ($n = 2$) or the smallest worthwhile change questions

($n = 3$) and could not be contacted to correct them. Figure 1 presents the frequencies of the smallest worthwhile change scores. On average, the improvement that participants would need to achieve to consider a self-management text message intervention worthwhile was 9.4 points (SD: 5.7; range 0–30), representing 31% of the total function score and 82% of participants' mean function score at baseline. 25%, 50% and 75% of the sample (25th percentile, median and 75th percentile) would need to achieve an improvement of at least 5.3, 9.0 and 12.0 points on the 0–30 scale, respectively, to consider the intervention to be worthwhile. These changes represent improvements of 18%, 30% and 40% of the total function score, respectively.

Table 1 Demographic and clinical characteristics of participants

Participants' characteristics, No. (%)	All participants ($n = 212$)
Sociodemographic characteristics	
Age, mean (SD) [IQR], y	59.8 (13.2) [51.0–59.0]
BMI, mean (SD) [IQR] kg/m^2	27.6 (5.9) [17.9–27.7]
Female sex	120 (56.6%)
Educational level	
Senior high school or below	30 (14.1%)
Certificate III/IV, diploma/advance diploma	55 (26.0%)
Bachelor's degree	50 (23.6%)
Post-graduate diploma/graduate certificate	31 (14.6%)
Masters, PhD or other	46 (21.7%)
Employment status	
Unemployed	15 (7.0%)
Part-time	43 (20.3%)
Full-time	58 (27.3%)
Volunteer	19 (9.0%)
Studying	4 (1.9%)
Full-time carer	8 (3.8%)
Retired	57 (26.9%)
Other	8 (3.8%)
General health	
Number of comorbidities, median [IQR]	1.0 [1.0–3.0]
Number of participants reporting sleep issues*	111 (52.3%)
Number of participants meeting PA guidelines (≥ 250 min mod-vig PA/week)	92 (43.4%)
Sedentary behaviour (hrs/week), mean (SD) [range]	8.4 (3.3) [6.0–10.1]
Quality of life (EQ-5L-5D) index, mean (SD) [range]	0.7 (0.1) [0.7–0.8]
Quality of life (EQ-5L-5D) VAS (0–100), mean (SD) [range]	64.0 (18.1) [50.0–80.0]
Low back pain clinical profile	
Duration of current episode, mean (SD) [IQR], weeks	7.4 (3.4) [4.0–11.0]
First episode of low back pain	34 (14.9%)
Presence of leg pain	98 (46.2%)
Care seeking in the past month	102 (48.1%)
Function (PSFS; 0–30), mean (SD) [IQR]	11.8 (5.4) [8.0–16.0]
Pain intensity (VAS; 0–100), mean (SD) [IQR]	51.5 (19.5) [35.2–67.5]

SD Standard deviation, *IQR* Interquartile range, *PA* Physical activity, *mod-vig* Moderate or vigorous

*Self-reported difficulties to fall asleep or waking up at night

Fig. 1 Distribution of the participant's reported smallest worthwhile change
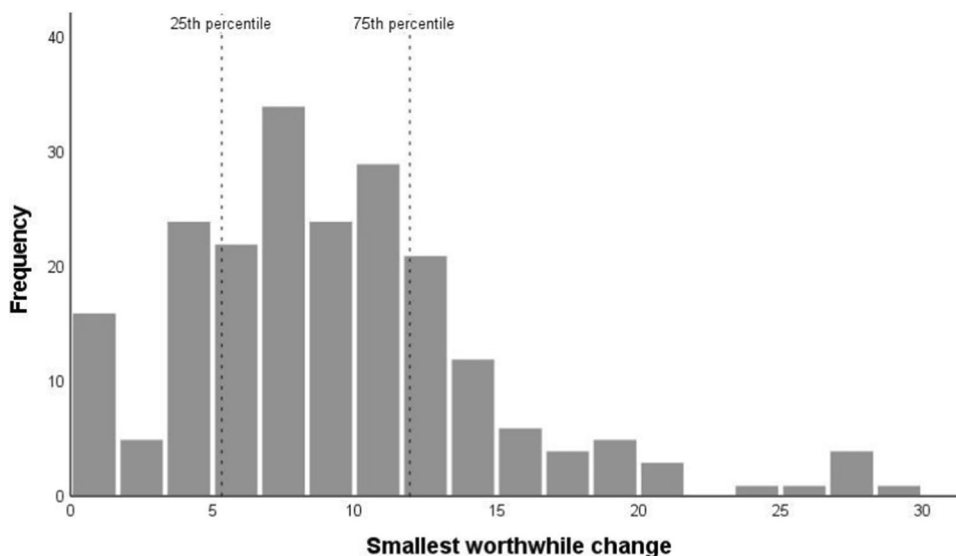
Table 2 Regression coefficients (95%CI; *p* value) of predictors of the multiple linear model for the smallest worthwhile change estimate

Predictors	Regression coefficient (95%CI)	*p* value
Age	0.01 (−0.03, 0.06)	0.571
Gender	0.30 (−1.01, 1.68)	0.630
Education	−0.20 (−1.48, 1.45)	0.982
Comorbidities	0.16 (−0.03, 0.06)	0.547
Sleep issues	−0.33 (−1.80, 1.14)	0.658
Sedentary behaviour	−0.05 (−0.22, 0.12)	0.544
Pain intensity	0.03 (−0.02, 0.07)	0.216
Function	−0.60 (−0.76, −0.44)	<0.001
Leg pain	0.41 (−0.91, 1.14)	0.539
Pain duration	−0.08 (−0.29, 0.13)	0.448
Quality of life EQ-5D index	3.77 (−1.36, 8.90)	0.149

Table 3 Regression coefficients (95%CI; *p* value) of predictors of the multiple linear model for the smallest worthwhile change estimate without outliers

Predictors	Regression coefficient (95%CI)	*p* value
Age	0.00 (−0.05, 0.04)	0.980
Gender	0.76 (−1.05, 1.21)	0.894
Education	0.92 (−0.36, 2.20)	0.157
Comorbidities	0.23 (−0.17, 0.63)	0.255
Sleep issues	−0.36 (−1.59, 0.86)	0.561
Sedentary behaviour	−0.02 (−0.18, 0.14)	0.800
Pain intensity	0.03 (−0.03, 0.07)	0.068
Function	−0.50 (−0.61, −0.38)	<0.001
Leg pain	0.38 (−0.77, 1.53)	0.513
Pain duration	−0.05 (−0.22, 0.12)	0.546
Quality of life EQ-5D index	3.72 (−0.34, 7.77)	0.072

The results of the multivariate regression model are presented in Table 2. Only baseline function was significantly associated with the elicited magnitude of the smallest worthwhile change. For each point decrease in function, there was an increase of 0.6 point in the smallest worthwhile change estimate ($b = -0.60$, 95% confidence interval [95% CI] −0.76, −0.44, $p < 0.001$). This effect size was medium ($\eta 2 = 0.219$). The results of the modified BP test indicated non-constant error variance ($\chi 2 (1) = 32.01$, $p < 0.001$) and robust standard errors were used. The normal Q–Q plot of standardised residuals and the Kolmogorov–Smirnov test indicated non-normal residuals (KS statistic (212) = 0.064, $p = 0.036$). Furthermore, outliers were detected, and a secondary analysis was conducted excluding the outlier observations. Results of the model effects and the associated 95% CI estimates based on robust standard errors are presented in

Table 3. The normal Q–Q plot and results of the Kolmogorov–Smirnov test indicated normality of residuals [KS statistic (201) = 0.058, $p = 0.094$]. Baseline function continued to be the only predictor associated with the smallest worthwhile change estimate. For each point decrease in function, there was an increase of 0.5 point in the smallest worthwhile change estimate ($b = -0.50$, 95% CI −0.61, −0.38, $p = <0.001$). The size of the effect of the baseline function score was large ($\eta 2 = 0.283$).

Discussion and conclusion

Discussion

The present study investigated the smallest change that people with non-specific, non-persistent low back pain would need to reach to consider a (text message-delivered) self-management intervention worthwhile given its costs, inconveniences and possible harms. For 50% of participants, an improvement of at least nine points (on a 0–30 point scale) was needed to make the intervention worthwhile, which represents an improvement of 30% of the total function scale score. Large variability in responses was observed. Of all predictors investigated, only function was associated with the magnitude of the smallest worthwhile change. People with worse function scores would need to see larger improvements in function to consider a self-management intervention worthwhile. Function scores explained 21.9% and 28.3% of the variance in the smallest worthwhile change estimate after accounting for the variance explained by other variables in the primary and secondary analysis, respectively.

Although function and disability are slightly different outcome measures, the current findings evidence that people with low back pain expect an improvement in function similar to the 30% improvement in disability expected with the natural course of the condition [22] to consider self-management worthwhile. However, it is important to note there is high variability in the estimates, showing that people would need to see vastly different changes in function to consider self-management worthwhile, from no change to full recovery. Interestingly, participants' characteristics did not explain this variability in the estimates other than their baseline function. Large variability was also reported by previous studies using the benefit-harm trade-off method to investigate the smallest worthwhile effect of interventions for low back pain [16, 17]. Given that the benefit-harm trade-off method holds all intervention's characteristics constant or undefined whilst only the effect of the intervention may change, researchers have argued that participants might value the undefined attributes differently, leading to the high variability in the estimates of the smallest worthwhile effect [17]. Since the current study applied a modified benefit-harm trade-off method, the same hypothesis could justify the variability found. Nonetheless, the reason why no association between participants' characteristics and the smallest worthwhile change estimate could be found was beyond the scope of the current study and limited comparisons with previous studies could be done.

This is the first study to investigate the smallest worthwhile change in function for people with low back pain and incorporate it in a clinical trial of low back pain. The estimates found may be used in future responder analysis by calculating differences in the proportions of people achieving their named threshold for the smallest worthwhile change between the intervention and control groups as well as the number needed to treat. This is a simple methodology, which has been shown to be feasible, not time-consuming for participants and could be easily incorporated into future trials, either added to online or printed baseline questionnaires. This could represent an interesting strategy to help elicit the clinical relevance of findings of primary outcomes of randomised controlled trials when used in responder analyses, especially in trials assessing the effectiveness of interventions on populations for which the smallest worthwhile effect is unknown.

Estimating the smallest worthwhile change at baseline surveys of randomised controlled trials presents some advantages over using anchor-based approaches (e.g. the minimal clinically important difference, or minimal important difference). The main advantages are (1) the definition of the smallest worthwhile change based on patients' perspectives and the threshold of what would be a minimal clinically relevant difference is not decided by researchers or based on clinimetric properties of the outcome measure, (2) estimates are intervention-specific and consider possible harms, inconveniences and costs of the intervention in question, and (3) the possibility of using the individualised estimates in a responder analysis. Inferences of relevant changes through anchor-based approaches might underestimate what is meaningful to patients. Previous studies have estimated the minimal important difference in function (also assessed through the Patient-Specific Functional Scale) for people with low back pain undergoing physiotherapy or educational and stretching sessions through anchor-based approaches [18–20]. They have shown that the minimal clinically important changes in this population lie between 0.8 and 1.3 points, representing 8% to 13% change in the total function score [18–20]. These findings are smaller than the estimates currently found (on average, people would need to achieve a 31% improvement in the total function score to consider self-management worthwhile). Furthermore, two studies have also estimated what would be medium and large clinically important differences in function, which would correspond to patients reporting being at least 'moderately better' and 'quite a bit better' on the global rating scale, respectively [20]. The medium and large changes in function would correspond to 13% and 43% improvements in the total function score [19, 20]. These estimates evidence that even when the clinically important differences are defined according to moderate improvements in the global rating scale (rather than small improvements) they might underestimate patients' perceptions. Thus, using the smallest worthwhile change rather than the anchor-based approaches estimates (e.g., the minimal clinically important difference)

in responder analysis consider patients' perspectives and can potentially lead to values closer to clinical practice.

Nonetheless, the current study has limitations that should be acknowledged. Although the sample was diverse and recruited from both the community and healthcare practices, it might not represent the perspectives of all the clinical population with low back pain. Participants could have been more motivated or have more time to engage in a self-management intervention than people with low back pain interested in other modalities of care since participants decided to enrol in the TEXT4myBACK trial (which is providing a self-management text message intervention). This might have led to smaller worthwhile change estimates limited to people with a non-persistent low back pain interested in this modality of care. Moreover, it is uncertain if people with persistent pain or radicular pain, for instance, would present similar estimates. It is also important to note that the smallest worthwhile change was estimated based on the Patient-Specific Functional Scale [10] and it is also uncertain if a similar estimate would be found if other function or disability outcome measures which assess other activities of daily living (such as the Oswestry Disability Index) were used. Additionally, certain attributes of the self-management intervention were kept from participants to ensure their blinding once randomised to the interventions, as the number of messages that would be sent, their frequency and time. It is possible that different thresholds would have been elicited if a more comprehensive description of the intervention had been provided. Given the estimates are intervention-specific, the smallest worthwhile change estimates found in the current study might not apply to different interventions involving other risks, inconveniences and costs. For instance, surgery might involve large costs, higher risks (such as infection) and inconveniences (such as hospitalisation, adverse events related to the procedure and time off work) and therefore, it is possible that patients would need to see greater improvements in their symptoms to consider the procedure worthwhile. Furthermore, the study had a cross-sectional design, it did not adopt a longitudinal perspective and a re-evaluation of participants' smallest worthwhile change after receiving the intervention. Although a previous study has evidenced that the estimates of the smallest worthwhile effects of anti-inflammatory medication and physiotherapy in people with persistent low back pain were similar when measured prior to treatment commencement and four weeks later, [17] we do not know if the smallest worthwhile change estimate found in the current study would remain the same if measured during or after the self-management intervention. Finally, the results of this study represent a worthwhile change in function over time, rather than an effect on function between groups. Thus, the current estimates should not be used to aid the interpretation of the clinical significance of effects found in randomised controlled trials and systematic reviews investigating self-management interventions.

Conclusion

People with non-specific, non-persistent low back pain reported that they need to improve nine points, on average, on a 0–30 function scale to consider a self-management intervention to be worthwhile. High variability was found between individual estimates (ranging from 0 to 30 points), highlighting the distinctive assessment made by each participant. However, there were no effects of demographic characteristics, comorbidities, lifestyle and low back pain-related factors on the magnitude of the estimate, except for function score. People with worse function scores require larger improvements to consider the intervention worthwhile.

Practice implications

The current estimates might be used in responder analyses of future randomised clinical trials investigating self-management interventions for low back pain. Alternatively, the estimates might also be used by clinicians to track patients' improvements when a self-management intervention is recommended. The methodology might be used by future randomised controlled trials when the intervention's smallest worthwhile effect is unknown. Similarly, this methodology might be used in prospective cohort studies to assess the clinical relevance of patients' improvements. Participants might be questioned at the start of the prospective study and after the intervention is received to assess if the changes in their symptoms were worthwhile considering the risks, inconveniences and harms associated with the treatment received.

Declarations

Conflict of interest The authors do not have any conflicts of interests to declare.

Ethical approval All procedures performed were in accordance with the ethical standards of the Australian National Statements on Ethical Conduct in Human Research and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The TEXT-

4myBACK randomised clinical trial was approved by the Northern Sydney Local Health District Human Ethics Committee (ETH 13895).

References

1. Herbert RD (2000) How to estimate treatment effects from reports of clinical trials. I: continuous outcomes. J Physiother 46:229–235. https://doi.org/10.1016/s0004-9514(14)60334-2
2. Ferreira ML, Herbert RD, Ferreira PH, Latimer J, Ostelo RW, Nascimento DP, Smeets RJ (2012) A critical review of methods used to determine the smallest worthwhile effect of interventions for low back pain. J Clin Epidemiol 65:253–261. https://doi.org/10.1016/j.jclinepi.2011.06.018
3. Ferreira M (2018) Research note: the smallest worthwhile effect of a health intervention. J Physiother 64:272–274. https://doi.org/10.1016/j.jphys.2018.07.008
4. McGlothlin AE, Lewis RJ (2014) minimal clinically important difference: defining what really matters to patients. JAMA 312:1342–1343. https://doi.org/10.1001/jama.2014.13128
5. Barrett B, Brown D, Mundt M, Brown R (2005) Sufficiently important difference: expanding the framework of clinical significance. Med Decis Making 25:250–261. https://doi.org/10.1177/0272989X05276863
6. Friedel JE, Foreman AM, Wirth O (2022) An introduction to "discrete choice experiments" for behavior analysts. Behav Process 198:104628. https://doi.org/10.1016/j.beproc.2022.104628
7. Fritsch CG, Ferreira PH, Prior JL, Clavisi O, Chow CK, Redfern J, Thiagalingam A, Lung T, McLachlan AJ, Ferreira ML (2021) TEXT4myBACK: a text message intervention to improve function in people with low back pain-protocol of a randomized controlled trial. Phys Ther 101(7):100. https://doi.org/10.1093/ptj/pzab100
8. Ware J Jr, Kosinski M, Keller SD (1996) A 12-Item short-form health survey: construction of scales and preliminary tests of reliability and validity. Med Care 34:220–233. https://doi.org/10.1097/00005650-199603000-00003
9. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, McLeod L, Delacqua G, Delacqua F, Kirby J, Duda SN (2019) The REDCap consortium: building an international community of software platform partners. J Biomed Inform 95:103208. https://doi.org/10.1016/j.jbi.2019.103208
10. Abbott JH, Schmitt JS (2014) The patient-specific functional scale was valid for group-level change comparisons and between-group discrimination. J Clin Epidemiol 67:681–688. https://doi.org/10.1016/j.jclinepi.2013.11.002
11. Sangha O, Stucki G, Liang MH, Fossel AH, Katz JN (2003) The self-administered comorbidity questionnaire: a new method to assess comorbidity for clinical and health services research. Arthritis Rheum 49:156–163. https://doi.org/10.1002/art.10993
12. Rosenberg DE, Norman GJ, Wagner N, Patrick K, Calfas KJ, Sallis JF (2010) Reliability and validity of the sedentary behavior questionnaire (SBQ) for adults. J Phys Act Health 7:697–705. https://doi.org/10.1123/jpah.7.6.697
13. Hawker GA, Mian S, Kendzerska T, French M (2011) Measures of adult pain: visual analog scale for pain (VAS Pain), numeric rating scale for pain (NRS Pain), McGill pain questionnaire (MPQ), short-form McGill pain questionnaire (SF-MPQ), chronic pain grade scale (CPGS), short form-36 bodily pain scale (SF-36 BPS), and measure of intermittent and constant osteoarthritis pain (ICOAP). Arthritis Care Res 63(Suppl 11):S240-252. https://doi.org/10.1002/acr.20543
14. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X (2011) Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 20:1727–1736. https://doi.org/10.1007/s11136-011-9903-x
15. Cohen J, Cohen P, West SG, Aiken LS (2002) Applied multiple regression/correlation analysis for the behavioral sciences. Routledge, New York
16. Christiansen DH, de Vos Andersen NB, Poulsen PH, Ostelo RW (2018) The smallest worthwhile effect of primary care physiotherapy did not differ across musculoskeletal pain sites. J Clin Epidemiol 101:44–52. https://doi.org/10.1016/j.jclinepi.2018.05.019
17. Ferreira ML, Herbert RD, Ferreira PH, Latimer J, Ostelo RW, Grotle M, Barrett B (2013) The smallest worthwhile effect of non-steroidal anti-inflammatory drugs and physiotherapy for chronic low back pain: a benefit-harm trade-off study. J Clin Epidemiol 66:1397–1404. https://doi.org/10.1016/j.jclinepi.2013.02.018
18. Maughan EF, Lewis JS (2010) Outcome measures in chronic low back pain. Eur Spine J 19:1484–1494. https://doi.org/10.1007/s00586-010-1353-6
19. Abbott JH, Schmitt J (2014) Minimum important differences for the patient-specific functional scale, 4 region-specific outcome measures, and the numeric pain rating scale. J Orthop Sports Phys Ther 44:560–564. https://doi.org/10.2519/jospt.2014.5248
20. Stratford P, Gill C, Westaway M, Binkley J (1995) Assessing disability and change on individual patients: a report of a patient specific measure. Physiother canada 47:258–263. https://doi.org/10.3138/ptc.47.4.258