SPINE An International Journal for the study of the spine, Publish Ahead of Print

DOI: 10.1097/BRS.0000000000003760

Does duration of pain at baseline influence longer-term clinical outcomes of low back pain patients managed on an evidence-based pathway?

Mary-Anne Jess [1] MSc, Cormac Ryan [1] PhD, Sharon Hamilton [1] PhD, Shaun Wellburn [1] PhD, Greg Atkinson [1] PhD, Charles Greenough [2] PhD, Glynis Peat [2] Dip, Andrew Coxon [2] PhD, Francis Fatoye [3] PhD, Diarmaid Ferguson [4,5] BSc, Alastair Dickson [5, 6] MRCGP, MB BS, LLM, MSc, BSc (Jt Hons), Helen Ridley[7] Dip, Denis Martin [1] DPhil

[1] School of Health and Life Sciences. Teesside University. Middlesbrough.

[2] South Tees Hospitals NHS Foundation Trust. Middlesbrough.

[3] Department of Health Professions, Faculty of Health Psychology and Social Care. Manchester Metropolitan University.

[4] Northumbria Healthcare NHS Trust, North Tyneside General Hospital, North Shields.

[5] Primary Care Rheumatology & Musculoskeletal Medicine Society.

[6] GP Lead for the North of England Back and Radicular Pain Pathway.

[7] Getting It Right First Time (GIRFT): Clinically-led improvement programme delivered in partnership with the Royal National Orthopaedic Hospitals NHS Trust, NHS England and NHS Improvement

Corresponding Author:

Mary-Anne Jess

School of Health and Social Care, Teesside University

Middlesbrough. TS1 3BX

07964234887

m.jess@tees.ac.uk

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

The Health Foundation Scaling Up Improvement Programme, the North East Academic Health Science Network and the Primary Care Rheumatology Society, supported the NERBPP.

Relevant financial activities outside the submitted work: consultancy, grants, patents.

Mini Abstract:

Following management on the NERBPP patients with shorter baseline duration of LBP demonstrated clinically important improvements on a suite of clinical outcome measures compared to those with ≥12 months duration of LBP. The latter may need additional support to achieve clinically relevant functional improvements in the medium-to-long term.

Study Design: Non-randomised longitudinal observational study.

Objective: To evaluate the association between baseline pain duration and medium-to-long term clinical outcomes, in low back pain (LBP) patients enrolled on the North East of England Regional Back Pain and Radicular Pain Pathway (NERBPP).

Summary of Background Data: The NERBPP is based upon National Institute for Health and Care Excellence (NICE) guidelines. These guidelines no longer differentiate management of LBP patients based on pain duration. Medium-to-long term data from the NERBPP is lacking.

Methods: Between May 2015 and December 2019, 786 and 552 LBP patients from the NERBPP returned 6-month and 12-month follow-up outcome measures, respectively. Outcomes included pain (Numerical rating scale), function (Oswestry Disability Index) and quality-of-life (EuroQol five-dimension, five-level questionnaire), analysed using a series of covariate-adjusted models. Patients were categorized into four groups based upon baseline pain duration: <3 months, ≥3 to <6 months, ≥6 months to <12 months, ≥12 months.

Results: Patients with <3 months duration demonstrated clinically important improvements on all outcomes, at both follow-ups. The improvements in outcomes from this group were larger than those in the ≥12 month's duration group ($p<0.05$), these group differences in change, in some cases surpassed our threshold for clinical relevance. Functional improvements in those with ≥12 month's duration were not clinically relevant at either follow-up. All patients, regardless of baseline pain duration, reported similar levels of readiness to self-manage at the 12-month follow-up.

Conclusions: Baseline pain duration would appear to be of clinical importance. Patients with shorter baseline pain duration demonstrated better outcomes. Those with ≥12 month's duration of pain may need additional support during their management to achieve clinically relevant functional improvements in the medium-to-long term. These findings raise questions about the decision by NICE to move away from duration of pain to differentiate management of LBP patients.

Key Words: low back pain, duration, disability, NICE guidelines, longer-term follow-up

Level of Evidence: 3

Introduction

Globally, the impact of low back pain (LBP) is increasing. It is the leading cause of disability, accountable for approximately 60.1 million years lived with disability in 2015, a 17.2% increase since 2005.[1] LBP results in high healthcare and societal costs.[2,3] Although most LBP episodes recover considerably in the first six weeks,[4] 65% of patients still report some pain after one year.[5]

Defining acute and chronic pain using duration is surrounded by debate as to where the most appropriate time-points lie to differentiate them,[6-8] and the suitability of using duration alone to show the dichotomy.[9,10] Whilst some clinical guidelines differentiate the management of LBP patients based on the traditional duration-based classification system (acute, subacute and chronic),[11,12] the National Institute for Health and Care Excellence (NICE) LBP guidelines (2016) instead recommend a management approach based on the use of risk stratification to classify patients.[13] This move away from a duration-based classification system has little, somewhat conflicting, empirical evidence supporting this shift in clinical practice. Dunn and Croft[14] found that patients with a longer duration of pain at baseline (≥3 years) were associated with poorer clinical outcomes. This work considered patient outcomes following a broad battery of usual care from their General Practitioner (GP). In contrast, recent work published by our group established that regardless of pain duration, all LBP participants demonstrated clinically relevant short-term improvements, on a suite of outcome measures, when managed on the evidence-based North East of England Regional Back Pain and Radicular Pain Pathway (NERBPP).[15] This clinical pathway was introduced in 2015 to standardise the management of LBP, based on the NICE LBP guidelines (2009).

It is known that LBP is characterised by variability, reoccurrence and remission,[16-19] therefore, considering longer-term outcomes would give a better clinical picture. The aim of this study was to extend prior work to explore the association between baseline pain duration and patient reported outcome measures (PROMS) in the medium-to-long-term, of LBP patients enrolled on the NERBPP.

Methods

Study design

This was a non-randomised, longitudinal, observational study, which was part of a large-scale evaluation of the implementation of the NERBPP.[20] This study was given ethical approval from Teesside University (Reference number R179/15). It used data gathered from South Tees Hospitals NHS Foundation Trust, one of the first adopter sites of the NERBPP.

In total 21,091 adults with LBP were referred onto the pathway by their GP between May 2015 and December 2019. This study focused on individuals who provided PROMS at the 6-month (n=786) and/or 12-month follow-ups (n=552).

The NERBPP was originally intended for acute LBP patients; nonetheless, patients of varying pain duration were referred onto the pathway. Patients were screened by their GP, using the STarT Back stratification tool.[21] Patients with low-risk of poor outcome on the

STarT Back tool were given advice on how to self-manage their symptoms and discharged. Patients with medium to high-risk of poor outcome were referred to a triage and treat practitioner (T&TP). The T&TP assessed all participants and referred them for the appropriate intervention; this included investigations and/or core therapies (physiotherapy incorporating education, exercise and/or manual therapy). For a small percentage of patients, who upon clinical assessment were adjudged to have a relatively higher risk of a poor outcome, a 100-hour residential, combined physical and psychological therapies program (CPPP) was offered. Participants included for analysis in this study may have received a combination of these management approaches.

Data collection

Baseline data was available on the STarT Back score, socio-demographic variables including age, sex, socioeconomic status, and a series of standardized, valid and reliable PROMS including pain, disability and quality-of-life; these were collected at the initial T&TP appointment. Other information obtained from the T&TP notes and uploaded to the electronic patient records system was the date of onset of the patient's symptoms and the date of GP referral onto the pathway. Baseline pain duration was calculated by subtracting the date of onset of symptoms, from date of referral onto the pathway. Patients were then categorized into four groups based upon their calculated pain duration: <3 months, ≥3 months to <6 months, ≥6 months to <12 months, ≥12 months.

At 6 and 12-months following discharge, participants were invited to complete follow-up outcome measures. The 11-point numerical rating scale (NRS) was used to measure pain.[22] Disability was assessed using the Oswestry Disability Index (ODI).[23] The EuroQol five-dimension, five-level (EQ-5D-5L) questionnaire was used to measure quality-of-life.[24] To determine the patient's readiness to self-manage, they were asked the question: "do you feel ready to self-manage your back pain?" The response was given using a 0-10-point continuous scale, modified from work by Lorig *et al.*, with 0 representing 'not confident' and 10 representing 'totally confident'.[25] For their overall perception of improvement the Global Subjective Outcome Scale (GSOS) was used, a six-point Likert-based scale with descriptors ranging from 'completely better' to 'worse'.[26] Participants were asked to describe their satisfaction with the service using the NHS Friends and Family Test (FFT); a six-point scale ranging from 'extremely likely' to 'extremely unlikely'.[27]

Analysis

Statistical analyses were carried out using SPSS (version 26). The data were manually screened for any data entry errors prior to commencing statistical analysis.

Initial analysis explored the PROMS of participants that provided outcome data following management on the NERBPP. Follow-up outcome data comprised the baseline data and 6-month and/or 12-month follow-up PROMS. Some participants that provided outcome data did not have a discharge code, therefore, may still have been on the pathway. However, most participants that provided outcome data were either discharged at their initial appointment

(same-day discharge (SDD)), discharged following an initial and at least one further appointment (standard discharge (StD)), while a small number were discharged due to non-attendance (non-attender (NA)), yet still provided follow-up outcome data.

Comparison of the outcomes for the four duration categories was undertaken using a series of covariate-adjusted models. Change scores for pain, disability and general health status were calculated by subtracting initial scores from follow-up scores (6-month and/or 12-month data). Duration category was the independent variable and the following were included as covariates: age, sex, socioeconomic status and baseline scores.[28] Minimal clinically important difference (MCID) for mean change scores were defined using recommendations by NICE (2016) guidelines as 10% improvement for continuous outcomes and 0.03 for EQ-5D. [13] A sensitivity analysis excluding non-attenders from the main analysis was performed to compare the influence they had on the mean change scores between the groups.

Data were presented as mean (standard deviation) for continuous measures while categorical data were presented as percentage, mean (standard deviation) or median (interquartile range). For categorical data the Kruskal-Wallis H test and/or Chi Square test was used. Statistical significance was set at $p< 0.05$.

Results

Of the 21,091 participants on the database, 12,685 participants were discharged and/or provided follow-up data. Of those 12,685 participants, 786 provided follow-up data (complete cases). Of those 786 participants, 425 were in the StD group, 131 were in the SDD group, 176 did not have a discharge code and 54 were in the NA group. The participant characteristics and baseline PROMS of participants that provided follow-up outcome data (n=786), compared to those that were discharged from the pathway but had not provided outcome data (incomplete cases, n=11,899), are shown in Supplementary Table A, http://links.lww.com/BRS/B657. Although there were statistically poorer baseline outcome measures for the incomplete cases, these differences were small and not clinically relevant. The baseline values for the PROMS for the pain duration groups are shown in Supplementary Table B, http://links.lww.com/BRS/B657 and Table C, http://links.lww.com/BRS/B657. There were no statistically or clinically significant differences at baseline between groups at the 6-month follow-up. The baseline difference in EQ-5D VAS for the 12-month follow-up data set, while statistically significant was small and not clinically relevant. Because any association between our proposed predictor of response (duration of pain) and baseline pain is important in the context of casual inference,[29,30] we used an equivalence testing approach to verify the lack of association in our data.[31] The largest mean difference in pain ratings (0-10 scale) between pain duration groups <3 months and ≥12 months was found to be 0.3 units, equating to a standardised mean difference of 0.15. The 90% confidence interval for the difference was -0.16 to 0.76 units. Both the lower and upper limits of this interval are below the 10% (1 unit) threshold we set for clinically relevant importance.

Fourteen of the participants that provided 6-month follow-up and 10 of those that provided 12-month follow-up data underwent the CPPP component of the pathway. All of these were in the ≥12 month's duration of pain group. To explore the impact of this, sensitivity analysis was undertaken with the data for these participants removed, there was no material change in the results (Supplementary Table D, http://links.lww.com/BRS/B657).

Outcome

The mean changes in outcome measures when grouping participants into one of the four pain duration categories from <3 months to ≥12 months is shown for the 6-month follow-up outcome data (n=786) in Table 1 and for 12-month outcome data (n=552) in Table 2.

At the 6-month follow-up, on average, all four groups improved on all PROMS. For those with <3 months duration, improvements were all clinically relevant. For those with longer durations, improvements were clinically relevant for pain and EQ-5D-Value but not for the ODI or EQ-5D VAS. There was a statistical difference between the groups, on all outcomes except EQ-5D VAS, with the general pattern of those with shorter pain durations improving the most. The difference between the shortest and longest durations were bordering on clinically significant. A similar pattern was demonstrated for the 12-month follow-up data, though the differences between the shortest and longest duration groups were slightly more pronounced; this was most evident for the ODI where the shortest duration group improved by a clinically greater amount than the longest duration group. Regarding the participants' perceptions of their ability to self-manage, there was little difference between the four groups at either follow-up point. Findings from both sensitivity analyses excluding non-attenders and the 14 CPPP participants were not materially different to the primary analysis (Supplementary Tables D, http://links.lww.com/BRS/B657 and E, http://links.lww.com/BRS/B657).

There was a significant difference between the groups on the GSOS, those with shorter pain durations reporting greater improvements at both follow-ups. Regarding recommending the service to a friend or relative there was no statistical difference between the duration groups at either follow-up (Tables 3 and 4).

Discussion

This analysis illustrates that LBP patients enrolled on the NERBPP improved at the 6-month and 12-month follow-ups, on all outcomes. Patients with <3 months baseline duration demonstrated improvements that were above the MCID recommended by NICE (2016) on all PROMS, at both follow-ups. The improvements in outcomes from this group were larger than those in the ≥12 month's duration group ($p<0.05$), and in some instances these group differences in change surpassed our threshold for clinical relevance. Although those with ≥12 month's duration showed clinically relevant improvements in pain and quality-of-life scores, functional improvements were not above the MCID at either time-point. The patients' GSOS scores reflected their clinical outcomes, with a greater proportion of those with <3 months duration reporting significantly better perception of improvements at both follow-ups. At the

12-month follow-up, regardless of baseline pain duration, all patients reported feeling some readiness to self-manage their symptoms; an important goal of the pathway. Patients reported similar levels of satisfaction with the service, regardless of baseline pain duration.

The NICE (2016) guidelines no longer differentiate the management of LBP patients based on pain duration, endorsing instead a risk stratification approach to classify patients. Initial findings previously published by our group,[15] appraising the outcomes on discharge from the NERBPP, found that all LBP patients, regardless of baseline pain duration, improved by clinically relevant amounts on a suite of outcomes. However, longer-term findings presented here suggest that baseline pain duration is of clinical importance. Those with ≥12 month's duration may need additional support to achieve clinically relevant functional improvements. These findings question the decision by NICE to move away from duration to differentiate the management of patients.

Our findings are consistent with previous research. An empirical study demonstrating similar associations, found that LBP patients with ≥3 years of pain were associated with worse clinical outcomes.[14] Additionally, a recent study examining the correlation between beliefs of staying active and function found that LBP patients with shorter pain durations were, statistically, significantly associated with clinically relevant functional outcomes.[32] Our findings are supported by a systematic review on the clinical course of LBP which highlights how patients with acute (defined as <12-weeks) or persistent symptoms improve markedly in the first six weeks; however, after this point patients with persistent LBP had lower improvements and could expect to have moderate levels of pain and disability at 12 months.[4] Another systematic review with comparable findings, while only focusing on the course of acute LBP, found that pain and disability improved rapidly within the first month, however over time improvements reduced and outcomes remained relatively constant.[33]

The main strength of this study is that data was collected as part of everyday clinical practice following management on the NERBPP, a pathway compliant with national guidelines. This increases generalizability of the findings in the real world to clinical practice in the UK, where the NICE guidelines are being implemented nationally through the National Back Pain Pathway of which the NERBPP was a forerunner. Another strength is that the PROMS included were valid, reliable and incorporated those recently recommended as a core outcome set, to ensure standardization of results.[34] Although a small amount of follow-up outcome data was obtained from patients that had been discharged due to non-attendance, and those that had received additional management through CPPP, a sensitivity analysis showed that removing non-attenders and CPPP participants from the analysis made no material difference (Supplementary Table A, http://links.lww.com/BRS/B657 and D, http://links.lww.com/BRS/B657).

A potential for bias in these findings is the high non-response rate to outcome measures at both follow-ups. The literature on improving postal response rates recommends: repeat mailing, telephone reminders and shorter questionnaires.[35-37] As this data was collected as part of routine care such strategies were not used. To investigate the potential bias of the high non-response rate, we compared participants' baseline characteristics for those that had completed follow-up outcome data (n=786), to those that had not (n=11899) and there was

little difference between them (Supplementary Table E, http://links.lww.com/BRS/B657). Another possible limitation of this work is the lack of a standardized question to establish the duration of the episode of LBP. This could potentially have led to misclassification of patients' pain duration. [38] As in our previous study,[15] duration of symptoms was obtained from the patient's notes, without clarification as to whether the duration referred to the patient's first onset of LBP or the current episode.

In studies that lack a comparator group, any associations between baseline status and change in status should be considered carefully.[29] In our study, the primary predictor of interest was duration of pain at baseline and not the rating of pain itself at baseline used in the calculation of change. It has been shown that an association between the proposed baseline predictor of response and the baseline value used in the calculation of change can bias inferences.[30] We found no association between the duration, and degree of pain at baseline (Supplementary Tables B, http://links.lww.com/BRS/B657 and C, http://links.lww.com/BRS/B657), and both baseline duration and status were entered into our statistical modal in an attempt to quantify independently adjusted influences. Nevertheless, we highlight the fact that our study was not a randomised controlled trial and so robust inferences about causality cannot be made.

Future research should explore if there are differences in other outcomes based upon baseline pain duration, such as return-to-work or surgery rates. Qualitative research, involving both patients and clinicians, could be undertaken to investigate what additional support those with longer-term pain would value.

Conclusion

This study found that patients with shorter durations of pain have superior outcomes following management on the NERBPP in the medium-to-long term compared to those with longer baseline pain duration. Baseline pain duration would appear to be of clinical importance and those with ≥12 months duration of pain may need additional support during their management to achieve clinically relevant functional improvements. These findings raise questions about the decision by NICE (2016) to move away from pain duration to differentiate management of LBP patients.

Acknowledgement:

The authors thank all staff, including Kevin Pears and Paul Green, and all the patients involved in the NERBPP.

Key Points:

- The NERBPP is an evidence-based clinical pathway for LBP and a forerunner to the National Back Pain Pathway.
- Following management on the NERBPP, patients with <3 months duration demonstrated improvements greater than the minimum clinically important difference (MCID) recommended by NICE (2016) on all clinical outcomes, at the 6 and 12-month follow-ups.
- Patients with ≥12 months duration displayed clinically relevant improvements in pain and quality of life, however, functional improvements were not above the MCID at either follow-up.
- Baseline pain duration would appear to be of clinical importance and those with ≥12 months duration of pain may need additional support during their management to achieve clinically relevant functional improvements in the medium-to-long term.

References

1. Kassebaum NJ, Arora M, Barber RM, et al. Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet. 2016;388(10053):1603-58.

2. Lin C-WC, Li Q, Williams CM, et al. The economic burden of guideline-recommended first line care for acute low back pain. European Spine Journal. 2018;27(1):109-16.

3. Geurts JW, Willems PC, Kallewaard J-W, et al. The impact of chronic discogenic low back pain: costs and patients' burden. Pain Research and Management. 2018;2018.

4. Costa LdCM, Maher CG, Hancock MJ, et al. The prognosis of acute and persistent low-back pain: a meta-analysis. Cmaj. 2012;184(11):E613-E24.

5. Itz CJ, Geurts J, Van Kleef M, et al. Clinical course of non□specific low back pain: A systematic review of prospective cohort studies set in primary care. European journal of pain. 2013;17(1):5-15.

6. Meucci RD, Fassa AG, Faria NMX. Prevalence of chronic low back pain: systematic review. Revista de saude publica. 2015;49:73.

7. Jess M-A, Hamilton S, Ryan CG, et al. Exploring the origin of low back pain sub-classification: a scoping review protocol. JBI database of systematic reviews and implementation reports. 2019;17(8):1600-6.

8. Jess M-A, Hamilton S, Ryan CG, et al. Exploring the origin of pain sub-classification, with emphasis on low back pain: a scoping review. JBI database of systematic reviews and implementation reports. In press.

9. Von Korff M, Dunn KM. Chronic pain reconsidered. Pain. 2008;138(2):267-76.

10. Cedraschi C, Robert J, Goerg D, et al. Is chronic non-specific low back pain chronic? Definitions of a problem and problems of a definition. The British Journal Of General Practice: The Journal Of The Royal College Of General Practitioners. 1999;49(442):358-62.

11. Qaseem A, Wilt TJ, McLean RM, et al. Noninvasive treatments for acute, subacute, and chronic low back pain: a clinical practice guideline from the American College of Physicians. Annals of internal medicine. 2017;166(7):514-30.

12. Stochkendahl MJ, Kjaer P, Hartvigsen J, et al. National Clinical Guidelines for non-surgical treatment of patients with recent onset low back pain or lumbar radiculopathy. European Spine Journal. 2018;27(1):60-75.

13. NICE. National Institute for Health and Care Excellence. Low back pain and sciatica in over 16s: assessment and management (NICE guideline NG59). Available from: https://www.nice.org.uk/guidance/ng59 2016 [11 December, 2019].

14. Dunn KM, Croft PR. The importance of symptom duration in determining prognosis. Pain. 2006;121(1-2):126-32.

15. Jess M-A, Ryan C, Hamilton S, et al. Does duration of pain at baseline influence clinical outcomes of low back pain patients managed on an evidence-based pathway? Spine. 2018;43(17):E998-E1004.

16. Croft PR, Macfarlane GJ, Papageorgiou AC, et al. Outcome of low back pain in general practice: a prospective study. Bmj. 1998;316(7141):1356.

17. Axén I, Leboeuf-Yde C. Trajectories of low back pain. Best Practice & Research Clinical Rheumatology. 2013;27(5):601-12.

18. Chen Y, Campbell P, Strauss VY, et al. Trajectories and predictors of the long-term course of low back pain: cohort study with 5-year follow-up. Pain. 2018;159(2):252.

19. Dunn KM, Jordan K, Croft PR. Characterizing the course of low back pain: a latent class analysis. American journal of epidemiology. 2006;163(8):754-61.

20. Martin D, Ryan C, Wellburn S, et al. The North of England Regional Back Pain and Radicular Pain Pathway: Evaluation. 2018.

21. Hill JC, Dunn KM, Lewis M, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. Arthritis Care & Research: Official Journal of the American College of Rheumatology. 2008;59(5):632-41.

22. Strong J, Ashton R, Chant D. Pain intensity measurement in chronic low back pain. The Clinical journal of pain. 1991;7(3):209-18.

23. Fairbank J, Couper J, Davies J, et al. The Oswestry low back pain disability questionnaire. Physiotherapy. 1980;66(8):271-3.

24. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of life research. 2011;20(10):1727-36.

25. Lorig K, Stewart A, Ritter P, et al. Outcome measures for health education and other health care interventions: Sage; 1996.

26. Harland N, Dawkin M, Martin D. Relative utility of a visual analogue scale vs a six-point Likert scale in the measurement of global subject outcome in patients with low back pain receiving physiotherapy. Physiotherapy. 2015;101(1):50-4.

27. NHS England. The Friends and Family Test. NHS England2014 [16 December, 2019]. Available from: https://www.england.nhs.uk/ourwork/pe/fft/.

28. Vickers AJ, Altman DG. Analysing controlled trials with baseline and follow-up measurements. Bmj. 2001;323(7321):1123-4.

29. Atkinson G, Batterham AM. True and false interindividual differences in the physiological response to an intervention. Experimental physiology. 2015;100(6):577-88.

30. Tennant PW, Arnold KF, Ellison GT, et al. Analyses of'change scores' do not estimate causal effects in observational data. arXiv preprint arXiv:190702764. 2019.

31. Lakens D. Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. Social psychological and personality science. 2017;8(4):355-62.

32. Riis A, Karran EL, Thomsen JL, et al. The association between believing staying active is beneficial and achieving a clinically relevant functional improvement after 52 weeks: a prospective cohort study of patients with chronic low back pain in secondary care. BMC Musculoskeletal Disorders. 2020;21(1):47.

33. Pengel LH, Herbert RD, Maher CG, et al. Acute low back pain: systematic review of its prognosis. Bmj. 2003;327(7410):323.

34. Chiarotto A, Deyo RA, Terwee CB, et al. Core outcome domains for clinical trials in non-specific low back pain. European Spine Journal. 2015;24(6):1127-42.

35. Nakash RA, Hutton JL, Jørstad-Stein EC, et al. Maximising response to postal questionnaires–a systematic review of randomised trials in health research. BMC medical research methodology. 2006;6(1):5.

36. Edwards PJ, Roberts I, Clarke MJ, et al. Methods to increase response to postal and electronic questionnaires. Cochrane database of systematic reviews. 2009(3).

37. Glidewell L, Thomas R, MacLennan G, et al. Do incentives, reminders or reduced burden improve healthcare professional response rates in postal questionnaires? Two randomised controlled trials. BMC health services research. 2012;12(1):250.

38. Dunn KM, de Vet HC, Hooper H, et al. Measurement of back pain episode inception in questionnaires: a study combining quantitative and qualitative methods. Journal of Musculoskeletal Pain. 2006;14(1):29-37.

Table 1: Mean change for patient reported outcome measure for 6-month follow-up, categorisation based on duration of pain.

Variable	n	< 3 months	≥3-<6 months	≥6-<12 months	≥12 months	*p*-value
Pain NRS (0-10) *	332	-2.3 (-2.8, -1.7) a	-1.8 (-2.5, -1.1)	-1.2 (-2.0, -0.5) c	-1.4 (-1.9, -1.0)	0.048
ODI (0-100%) *	454	-14.1 (-17.0, -11.3) a	-6.8 (-10.4, -3.3) d	-11.1 (-15.2, -6.9)	-5.0 (-7.2, -2.8) z	<0.001
EQ-5D Value (1 to -0.594) #	458	0.19 (0.15, 0.23) a	0.15 (0.10, 0.19) b	0.12 (0.06, 0.18)	0.08 (0.05, 0.11)	0.001
EQ-5D VAS (0-100%) #	453	11.2 (7.6, 14.9) a	9.2 (4.7, 13.6)	7.3 (1.8, 12.7)	7.2 (4.3, 10.0)	0.346
Self-management #	417	5.6 (5.0, 6.1)	5.2 (4.5, 6.0)	4.2 (3.4, 4.9) c	5.5 (5.0, 5.9) z	0.021

Data are Mean change (95% Confidence Interval Lower, Upper Bound), Mean (SD) and Median (IQR) by use of covariate adjusted models for: age, sex, socioeconomic status and baseline score for the outcome measure. a: statistically significant better outcome at < 3 months than at ≥12 months. b: statistically significant better outcome at ≥3-<6 months than at ≥12 months. c: statistically significant better outcome at <3 months than at ≥6- <12 months. d: statistically significant better outcome at <3 months than at ≥3- <6 months. z: statistically significant better outcome at ≥6- <12 months than at ≥12 months. Not all participants provided data for each of the variables, numbers are given for each duration category (n= < 3 months; ≥3-<6 months; ≥6<12 months; ≥12 months) NRS= Numerical Rating Scale (n=97; 50; 42; 143). ODI= Oswestry Disability Index (n= 119; 79; 57; 199). EQ-5D= EuroQol five-Dimension Questionnaire, Value (n=121; 82; 54; 201) VAS= Visual Analogue Scale (n= 123; 82; 55; 198). Self-management (n= 113; 62; 58; 182)

*Higher NRS and ODI scores are worse.

#Lower EQ-5D Value, EQ-5D VAS and self-management scores are worse.

Table 2: Mean change for patient reported outcome measure for 12-month follow-up, categorisation based on duration of pain.

Variable	n	< 3 months	≥3-<6 months	≥6-<12 months	≥12 months	*p*-value
Pain NRS (0-10) *	280	-2.3 (-2.8, -1.7)	-2.0 (-2.7, -1.4)	-1.8 (-2.5, -1.0)	-1.6 (-2.1, -1.1)	0.388
ODI (0-100%) *	291	-15.5 (-19.5, -11.6) a	-10.7 (-15.4, -6.0) b	-10.2, (-15.1, -5.3) c	-3.5 (-6.7, -0.3) z	<0.001
EQ-5D Value (1 to -0.594) #	295	0.20 (0.15, 0.25) a	0.14 (0.08, 0.19) b	0.15 (0.09, 0.22)	0.07 (0.03, 0.12) z	0.002
EQ-5D VAS (0-100%) #	293	13.2 (8.5, 17.9) a	7.9 (2.4, 13.4)	8.0 (2.1, 13.9)	5.3 (1.4, 9.3)	0.094
Self-management #	338	6.0 (5.4, 6.6) a	4.9 (4.2, 5.6) d	4.9 (4.1, 5.6) c	5.2 (4.7, 5.7)	0.053

Data are Mean change (95% Confidence Interval Lower, Upper Bound), Mean (SD) and Median (IQR) by use of covariate adjusted models for: age, sex, socioeconomic status and baseline score for the outcome measure. a: statistically significant better outcome at < 3 months than at ≥12 months. b: statistically significant better outcome at ≥3-<6 months than at ≥12 months. c: statistically significant better outcome at <3 months than at ≥6- <12 months. d: statistically significant better outcome at <3 months than at ≥3- <6 months. z: statistically significant better outcome at ≥6- <12 months than at ≥12 months. Not all participants provided data for each of the variables, numbers are given for each duration category (n= < 3 months; ≥3-<6 months; ≥6<12 months; ≥12 months) NRS= Numerical Rating Scale (n=73; 53; 46; 108). ODI= Oswestry Disability Index (n= 75; 52; 49; 115). EQ-5D= EuroQol five-Dimension Questionnaire, Value (n=77; 58; 47; 113) VAS= Visual Analogue Scale (n= 76; 57; 49; 111). Self-management (n= 89; 62; 56; 133)

*Higher NRS and ODI scores are worse.

#Lower EQ-5D Value, EQ-5D VAS and self-management scores are worse.

Table 3: Categorical data for GSOS and FFT for 6-month follow-up, categorisation based on duration of pain.

Variable	< 3 months	≥3-<6 months	≥6-<12 months	≥12 months	*p*-value
GSOS	n=164	n=95	n=82	n=259	<0.01
Completely better	11%	4.2%	6.2%	1.5%	
A lot better	29.3%	25.3%	21%	20.1%	
Moderately better	11.6%	17.9%	4.9%	5.8%	
A little better	18.3%	16.8%	14.8%	19.3%	
Same	23.2%	27.4%	42 %	36.7%	
Worse	6.7%	8.4%	11.1%	16.6%	
FFT	n=90	n=62	n=56	n=132	0.124
Extremely likely	32.7%	27.4%	30.5%	29.5%	
Likely	32.7%	37.9%	26.8%	26.9%	
Neither likely or unlikely	17.6%	13.7%	22.0%	22.0%	
Unlikely	16.1%	12.6%	9.8%	8.0%	
Extremely unlikely	3.0%	5.3%	4.9%	9.5%	
Don't know	7.9%	3.2%	6.1%	4.2%	

GSOS= Global Subjective Outcome Scale. FFT= Friends and Family Test.

Table 4: Categorical data for GSOS and FFT for 12-month follow-up, categorisation based on duration of pain.

Variable	< 3 months	≥3-<6 months	≥6-<12 months	≥12 months	*p*-value
GSOS	n=90	n=62	n=56	n=132	0.034
Completely better	12.2%	6.5%	3.6%	3.8%	
A lot better	35.6%	29.0%	23.2%	18.2%	
Moderately better	7.8%	16.1%	14.3%	9.8%	
A little better	13.3%	14.5%	12.5%	12.1%	
Same	23.3%	24.2%	35.7 %	40.2%	
Worse	7.8%	9.7%	10.7%	15.9%	
FFT	n=90	n=62	n=56	n=132	0.381
Extremely likely	37.8%	22.6%	26.8%	22.7%	
Likely	21.1%	40.3%	28.6%	33.3%	
Neither likely or unlikely	17.8%	16.1%	19.6%	17.4%	
Unlikely	10.0%	9.7%	14.3%	11.4%	
Extremely unlikely	11.1%	11.3%	8.9%	9.8%	
Don't know	2.2%	0.0%	1.8%	5.3%	

GSOS= Global Subjective Outcome Scale. FFT= Friends and Family Test.