PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	SUcceSS: Surgery for Spinal Stenosis – protocol of a randomised, placebo-controlled trial
AUTHORS	Anderson, David; Ferreira, Manuela; Harris, Ian; Davis, Gavin; Stanford, Ralph; Beard, David; Li, Qiang; Jan, Stephen; Mobbs, Ralph J.; Maher, Christopher; Yong, Renata; Zammit, Tara; Latimer, Jane; Buchbinder, Rachelle

VERSION 1 – REVIEW

REVIEWER	Carlo Ammendolia University of Toronto Toronto, Canada
REVIEW RETURNED	09-Jul-2018

GENERAL COMMENTS	This study protocol addresses a very important, relevant and timely research question related to the effectiveness of decompression surgery compared to placebo surgery for neurogenic claudication due to lumbar spinal stenosis. The rationale for the study is well described and convincing, and proposed study design and methods are sound and of high scientific vigour (except for the item below). The main methodological concern relates to the two primary outcomes selected for this study. The ODI is most commonly used to assess outcomes in low back pain studies. The population of interest in this clinical trial is neurogenic claudication, which by definition, relates to lower extremity symptoms impacting walking ability. Individuals with neurogenic claudication often do not complain of low back pain. Walking impairment is the dominant functional limitation in this population (Ammendolia 2017). Improvements in walking ability and reductions in lower extremity pain are the most meaningful outcomes in patients with neurogenic claudication (Ammendolia 2017). The self-paced walk test is the only validated objective walking measure in this population and the ODI had been shown not to be an adequate surrogate measure of walking ability (see your reference #31, Tomkins-Lane et al 2014). The other co-primary outcome measure is the walking capacity change scale, which to my knowledge has not been validated in this population. Moreover, this outcome measure provides no information on the change in walking distance or degree of limitations. The authors may consider using the walking section of the ODI as a primary outcome (not secondary) since this measure adequately correlates to the self-paced walk test and provides a measure of walking limitations (Tomkins-Lane 2014). Other considerations include: Methods Participants and Recruitment

	1) Indicate the number of surgeons expected to participate in the study and expected recruitment dates. 2) How will you convince eligible potential participants to participate in the study when their surgeon has told them that they need decompression surgery and by agreeing to participate they will have a 50% chance of getting the recommended surgery? Are there incentives planned for patients participating in the study? Inclusion Criteria 1) Define failure to improve with non-surgical treatment. Would a single failed epidural injection qualify? Exclusion Criteria 1) Recommend to use of the ankle-brachial index to assess peripheral vascular disease rather than foot pulses (due high false negative and false positive) or Doppler ultrasound (need over 50% occlusion before detected (see Ammendolia JCCA 2014). 2) Would you exclude a patient with neurogenic claudication with lumbar stenosis who has underlying Parkinson's disease? It may be best to exclude patients with neurological conditions that can impact walking ability. Primary outcomes 1) See above comments Secondary Outcomes 1) Please indicate whether you will be using the combined scores of functional and symptoms scales of the Swiss Spinal Stenosis questionnaire or independent scores 2) It may be more meaningful to ask about the average lower extremity and low back pain (NPS) scores while walking/standing since these patients are often asymptomatic otherwise. 3) Please indicate the MCID for each secondary outcome Statistical analyses 1) Is an "as treated" analysis planned if there is a high cross over from placebo to actual decompression during longer-term follow-ups (> 3 months)? 2) If high cross over rates how will this impact the cost-effective analysis? References 1) Reference 30 and 31 are the same. General 1) Will participants be notified of the study results and their assigned allocation at the end of the study? 2) The checklist requires contact information of the sponsor be provided.

REVIEWER	Peter Försth Inst of surgical science Uppsala University Hospital Uppsala University Sweden
REVIEW RETURNED	11-Aug-2018

GENERAL COMMENTS	Much credit to the authors for the efforts in planning the SUcceSS trial! If the study can be performed as planned it will support spine surgeons and health care providers with the highest level of evidence for the efficacy of decompressive surgery for lumbar spinal stenosis. An ageing population and an increased desire among

	elderly to stay active will put strain on health care systems worldwide. Because the annual rate of surgery for the most commonly surgically treated spinal condition (i.e. LSS) have increased in the past decades, it is important to evaluate the efficacy of the surgical procedure. The placebo approach and the blinding arrangements rule out any bias in performance or expectancy. The lack of external validity that RCTs often are accused for is excellently handled with the large separate observational arm.

REVIEWER	Kazunari Fushimi Gifu University Graduate School of Medicine, Japan.
REVIEW RETURNED	15-Aug-2018

GENERAL COMMENTS	The authors report a protocol for a randomized placebo-controlled trial of decompressive surgery for symptomatic LSS. This study addresses an important topic in clinical medicine. All participants and assessors are blinded to treatment allocation. This kind of placebo-controlled trial is ethically challenging study, but the protocol has been approved by ethics committees. Ethical issue seems to be cleared. Data analysis is conducted by both randomized controlled trial analysis and intention-to-treat analysis. The authors use an appropriate outcome measures, such as pain analysis (NRS for low back pain and lower limb pain), functional status (ODI), and quality of life (QALYS). A few points should be considered: Major points: 1. Participants who decline to participate in RCT but consent to be observational cohort are followed–up (n=200), and finally, included in collected patients data (Figure 1). This point makes us confused, and difficult to understand. If these patients are included in the final data, this study is not double-blinded placebo-controlled trial any more. Please clarify. 2. If one participant was allocated in Placebo group, but he/she is not satisfied with the result of surgery and wish to receive proper decompressive surgery, how does the participant receive another surgery? And, how will the participant be evaluated in this study? 3. In decompression group, surgeon (unblinded person) should confirm whether decompression surgery is effectively completed including removal of osseous spur and ligmentum flavum by MRI or CT. If decompression surgery is not successful, placebo-comparative study is not reliable. How do authors evaluate the result of surgery. 4. Patients with central canal stenosis seem to be included as participants of the study. Will patients with foraminal stenosis (intra- or extra- foraminal pathology) be excluded from this study? How do authors distinguish them? It should be clarified. Minor points: 1. In Abstract, line 4, the word “ublind”, is the spell correct? 2. Author described that surgery is performed with a single midline or dual paramedian longitudinal skin incision. Which approach will be chosen for Placebo surgery? How does surgeon choose a way of approach?

REVIEWER	Nick Jeffery Texas A&M University

REVIEW RETURNED	18-Aug-2018

GENERAL COMMENTS	General comments Thank you for inviting me to review this interesting trial protocol. It is great to see a placebo controlled trial on this subject since – as the authors point out – this is a much more robust design to determine whether decompressive surgery is beneficial for lumbar spinal canal stenosis. Specific comments P9, line7: more cost-effective?? Odd claim for surgery versus placebo! Presumably authors mean that surgery would provide more benefit than the presumed conservative effect of placebo? P10, line 2: inclusion of claudication patients only may limit generalizability. Although does prove the point about these specific patients. P10, inclusions: 'have not improved' – is this no improvement at all …? It might help to clarify your meaning here. Exclusions: a) There is no psychological assessment included here - is that deliberate? If so, why? b) Might need better definition of CE syndrome – P10, line 35. What is your definition of CE syndrome and how severe wouid it have to be to be an exclusion? P11, blinding. Is this feasible in all the sites – to keep the blinding working? P13, line 38 Primary outcomes: it is a little concerning that the primary outcome timepoint will be at 3 months - this is very early to judge the outcome of surgery fairly against conservative (placebo) therapy. There is a possibility of a high number of unsatisfied patients at 3 months wanting to be unblinded with the hope of receiving 'true' surgery. This will then impact on long-term outcome assessment - which is really of more interest to patients than 3-month outcome. P12, line 38: There is no definition of 'poor' clinical outcome. This will need to be more clearly defined to avoid losing a high proportion of patients from blinded assessment at later timepoints. It would seem to be better to only break blinding for those that are doing very poorly (eg if they have since developed severe CE disease). P15: Would like to see more detail in adverse events – what sort of things, what will you do about them etc. P16, line 27: although this will be complicated by the patients' knowledge of treatment group – can argue should not be included. P16, line 37: A bit of a contentious statistical point - there are no true continuous outcomes here! – these are all ordered categories. P18, line 24: these are pretty big differences - change in ODI of 15 points in a 50 pint scale or between 30% and 60% recovery rate - so

	negative results may not be enough to definitively persuade surgeons that surgery is not worthwhile.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Carlo Ammendolia

1. The main methodological concern relates to the two primary outcomes selected for this study. The ODI is most commonly used to assess outcomes in low back pain studies. The population of interest in this clinical trial is neurogenic claudication, which by definition, relates to lower extremity symptoms impacting walking ability. Individuals with neurogenic claudication often do not complain of low back pain. Walking impairment is the dominant functional limitation in this population (Ammendolia 2017). Improvements in walking ability and reductions in lower extremity pain are the most meaningful outcomes in patients with neurogenic claudication (Ammendolia 2017). The self-paced walk test is the only validated objective walking measure in this population and the ODI had been shown not to be an adequate surrogate measure of walking ability (see your reference #31, Tomkins-Lane et al 2014).

We thank Dr Ammendolia for his review of our protocol. However, we have carefully considered the outcome domains and the best instruments to measure these (validity, reliability and responsiveness) in this trial, and we reached consensus among the expert trial group based upon these data for the co-primary outcomes we have chosen. Our selection of co-primary outcomes reflects the two main complaints of patients with lumbar spinal stenosis (LSS): walking capacity and function. Leg symptoms with walking (claudication) are relevant because of their impact on walking ability (i.e. the loss of function is more important than the level of pain when attempting to walk). Although we agree that the ODI is commonly used to assess function in low back pain studies, it is the preferred tool to measure function in randomised trials of surgery for LSS, and included in the main and largest surgical trials (e.g. Forsth et al 2016; Weinstein et al 2008; Delitto et al 2015).1-3 It has also been validated in this population, showing good responsiveness after surgery, excellent internal consistency (alpha Cronbach of 0.9) and test retest reliability (ICC:0.89).4 The ODI also correlates most strongly with patient satisfaction after surgery, compared to the Swiss Spinal Stenosis questionnaire and Shuttle Walking Test.4 We have clarified this in the manuscript (page 12, paragraphs 1 and 2):

Our selection of co-primary outcomes reflects the two main complaints of patients with lumbar spinal stenosis (LSS): walking capacity and function. The ODI is a self-reported questionnaire commonly used to assess function in people with spinal conditions, including lumbar spinal stenosis.1-3 The outcome has been validated in this population, showing good post-surgery responsiveness, excellent internal consistency (alpha Cronbach of 0.9) and test retest reliability (ICC:0.89).4 The ODI is also strongly correlated with patient satisfaction after surgery.4

2. The other co-primary outcome measure is the walking capacity change scale, which to my knowledge has not been validated in this population. Moreover, this outcome measure provides no information on the change in walking distance or degree of limitations. The authors may consider using the walking section of the ODI as a primary outcome (not secondary) since this measure adequately correlates to the self-paced walk test and provides a measure of walking limitations (Tomkins-Lane 2014).

The walking capacity measure has been validated in this population and showed strong correlation with the ODI walking section (Spearman Rho:0.64) and self-paced walking test (Spearman Rho:0.65;

ICC: 0.68).5 6 We have chosen to keep the ODI walking section as a secondary outcome, given it is already measured in the full ODI questionnaire – a co-primary outcome. This information has also been added to the manuscript (page 10):

The self-reported walking capacity measure has also been validated in people with LSS and neurogenic claudication, showing strong correlation with the ODI walking section (Spearman Rho:0.64) and self-paced walking test (Spearman Rho:0.65; ICC: 0.68).5 6

3. Indicate the number of surgeons expected to participate in the study and expected recruitment dates.

We have currently recruited 8 surgeons to participate in the study. This number is likely to increase as an additional 5 surgeons have demonstrated interest in being involved. Given we are in the recruitment stage of the trial, we believe this information will be more accurate and better placed in the trial main results report.

4. How will you convince eligible potential participants to participate in the study when their surgeon has told them that they need decompression surgery and by agreeing to participate they will have a 50% chance of getting the recommended surgery? Are there incentives planned for patients participating in the study?

As stated on page 10 of the manuscript, trial surgeons will verify eligibility and invite consecutive, eligible patients to be included in the study. By their involvement in the study and in the absence of high quality evidence, they have equipoise regarding the value of surgery and this will be conveyed to their patients. In addition, a patient information sheet disclosing the nature (i.e. 50% chance of being allocated to placebo surgery) and aims of the study will be given to and discussed with potential participants by trial staff, as recommended by GCP. We do not aim to 'convince' patients to participate and will not provide financial incentives. Our investigators have previously recruited successfully for placebo surgery trials based on a similar structure7.

5. Define failure to improve with non-surgical treatment. Would a single failed epidural injection qualify?

Following current and recommended clinical practice, patients who present to trial surgeons with symptomatic central canal lumbar spinal stenosis of at least 3-month duration and have not improved with non-surgical care (e.g. physiotherapy, medication, epidural injection) will be deemed eligible (provided all other inclusion criteria and exclusion criteria are met). This is stated on page 9 of the manuscript. In the absence of high quality evidence establishing the efficacy of non-surgical interventions for this condition, provided the patient has had symptoms for at least 3 months, a single failed epidural injection would qualify for failing to improve with non-surgical care.

6. Recommend to use of the ankle-brachial index to assess peripheral vascular disease rather than foot pulses (due high false negative and false positive) or Doppler ultrasound (need over 50% occlusion before detected (see Ammendolia JCCA 2014).

We thank Dr Ammendolia for the recommendation. Although the specificity of ABI in diagnosing peripheral vascular disease is high (ranging from 83% to 99%) its sensitivity can be as low as 15%, especially in older individuals8. We have chosen an approach that is routinely used by our participating surgeons, ensuring the pragmatic nature of our trial design. Therefore, patients will be excluded if they present with known or demonstrated peripheral vascular disease causing vascular claudication i.e. claudication accompanied by absent foot pulse or vascular insufficiency detected with Doppler Ultrasound or CT angiography (page 9 of the manuscript).

7. Would you exclude a patient with neurogenic claudication with lumbar stenosis who has underlying Parkinson's disease? It may be best to exclude patients with neurological conditions that can impact walking ability.

This trial intends to be a pragmatic, real-world trial, where experienced spine surgeons deem that a patient requires lumbar decompression, who is then recruited into the trial. As such, mild Parkinsonism is not a contra-indication to such surgery, and is assessed by the treating surgeon at initial consultation, provided that the patient fulfils all of the inclusion criteria and there are no concerns regarding full follow up. Similarly, other mild forms of neurodegenerative disease are not absolute contraindications.

8. Please indicate whether you will be using the combined scores of functional and symptoms scales of the Swiss Spinal Stenosis questionnaire or independent scores.

We will be presenting separate results for the symptom severity subscale, physical functional subscale and satisfaction subscale9. This has been clarified in the manuscript (page 13).

Results of the symptom severity and functional subscales of the Swiss Spinal Stenosis Questionnaire will be reported separately.

9. It may be more meaningful to ask about the average lower extremity and low back pain (NPS) scores while walking/standing since these patients are often asymptomatic otherwise.

We thank Dr Ammendolia for this suggestion. The NPS will be used to capture average pain severity in the past week, and will not exclude pain during walking or standing. Moreover, pain during walking will be captured in the other secondary outcomes, Spinal Stenosis questionnaire, the walking section of the ODI and the co-primary outcome (improvement in walking ability).

10. Please indicate the MCID for each secondary outcome.
The MCID for the continuous, secondary outcome measures are presented below. We would like to emphasise that these thresholds are of limited use when interpreting the results of randomised controlled trials, as previously shown by our group.10 11

Walking ability (ODI): not available.
Swiss Spinal Stenosis Questionnaire
- Physical function subscale: 0.52.
- Symptom severity subscale: 0.48.
Pain numerical scale: 1 point

11. Is an "as treated" analysis planned if there is a high cross over from placebo to actual decompression during longer-term follow-ups (> 3 months)?

Yes, although we do not expect high cross overs from placebo to decompression during longer term follow up (>3 months) we plan secondary 'per protocol' analyses as well as the primary intention to treat analysis.

12. If high cross over rates how will this impact the cost-effective analysis?

Please see answer above. Secondary cost-effectiveness analyses may be conducted based on the per protocol analysis reports.

13. Reference 30 and 31 are the same.

This has been corrected.

14. Will participants be notified of the study results and their assigned allocation at the end of the study?

At the end of the trial, the results of the study or treatment allocation may be provided to participants who choose to receive them and contact the research team. This has been clarified in the manuscript (page 16).

At the end of the trial, participants may be notified of the study results and treatment allocation if they wish to be.

15. The checklist requires contact information of the sponsor be provided.
The corresponding author is the contact person for the sponsor. This information has been added to the checklist.

Reviewer: 2
Reviewer Name: Peter Försth

1. Much credit to the authors for the efforts in planning the SUcceSS trial! If the study can be performed as planned it will support spine surgeons and health care providers with the highest level of evidence for the efficacy of decompressive surgery for lumbar spinal stenosis. An ageing population and an increased desire among elderly to stay active will put strain on health care systems worldwide. Because the annual rate of surgery for the most commonly surgically treated spinal condition (i.e. LSS) have increased in the past decades, it is important to evaluate the efficacy of the surgical procedure. The placebo approach and the blinding arrangements rule out any bias in performance or expectancy. The lack of external validity that RCTs often are accused for is excellently handled with the large separate observational arm.

We thank Dr Forsth for his positive comments.

Reviewer: 3
Reviewer Name: Kazunari Fushimi

1. The authors report a protocol for a randomized placebo-controlled trial of decompressive surgery for symptomatic LSS. This study addresses an important topic in clinical medicine. All participants and assessors are blinded to treatment allocation. This kind of placebo-controlled trial is ethically challenging study, but the protocol has been approved by ethics committees. Ethical issue seems to be cleared. Data analysis is conducted by both randomized controlled trial analysis and intention-to-treat analysis. The authors use an appropriate outcome measures, such as pain analysis (NRS for low back pain and lower limb pain), functional status (ODI), and quality of life (QALYS).

Participants who decline to participate in RCT but consent to be observational cohort are followed–up (n=200), and finally, included in collected patient data (Figure 1). This point makes us confused, and difficult to understand. If these patients are included in the final data, this study is not double-blinded placebo-controlled trial any more. Please clarify.

Thank you for your comments. Only the patients in the randomised cohort will be part of the effectiveness analysis. This has been clarified in the statistical analysis section, page 17:

Treatment effectiveness analyses of randomised trial data will be blinded and performed on an intention-to-treat basis.

2. If one participant was allocated in Placebo group, but he/she is not satisfied with the result of surgery and wish to receive proper decompressive surgery, how does the participant receive another surgery? And, how will the participant be evaluated in this study?

Participants will be encouraged to remain blinded following randomisation for at least 3 months (the primary endpoint). This is consistent with surgeons' current practice for considering re-operation, unless a serious adverse event (e.g. foot drop or cauda equina) presents, in which case the surgeon may operate immediately. After this point, the need for unblinding and further surgery will be at the patient's and surgeon's discretion. Data analysis will be performed on an intention-to-treat basis (page 15 of the manuscript).

3. In decompression group, surgeon (unblinded person) should confirm whether decompression surgery is effectively completed including removal of osseous spur and ligmentum flavum by MRI or CT. If decompression surgery is not successful, placebo-comparative study is not reliable. How do authors evaluate the result of surgery?

We will be performing an MRI on a random 20% sample of the randomised cohort at 2 years, to assess for treatment fidelity. Our un-blinded study member will also have access to surgeon's operating notes, which will provide detailed information of the surgical procedure. This information will be used to ensure treatment allocation has been completed. It should be noted that the thoroughness of decompression is judged intra-operatively, not with routine post-operative scanning.

4. Patients with central canal stenosis seem to be included as participants of the study. Will patients with foraminal stenosis (intra- or extra- foraminal pathology) be excluded from this study? How do authors distinguish them? It should be clarified.

Patients will need to have neurogenic claudication, and predominant central canal stenosis on imaging. However, most patients with central canal stenosis also have some foraminal stenosis and foraminal decompression is part of the laminectomy (decompression) procedure.

5. In Abstract, line 4, the word "ublind", is the spell correct?

The spelling in the abstract has been corrected.

6. Author described that surgery is performed with a single midline or dual paramedian longitudinal skin incision. Which approach will be chosen for Placebo surgery? How does surgeon choose a way of approach?

Surgeons will be free to choose their preferred skin incision. Patients will not be randomised until after their incision and randomisation is being stratified by surgeon, therefore type of skin incision should be similar across treatment groups.

Reviewer: 4
Reviewer Name: Nick Jeffery

1. Thank you for inviting me to review this interesting trial protocol. It is great to see a placebo controlled trial on this subject since – as the authors point out – this is a much more robust design to determine whether decompressive surgery is beneficial for lumbar spinal canal stenosis. Specific commentsP9, line7: more cost-effective?? Odd claim for surgery versus placebo! Presumably authors mean that surgery would provide more benefit than the presumed conservative effect of placebo?

Yes, as stated in the manuscript, page 18: In the event of an observed positive treatment effect, a cost-effectiveness analysis will be conducted.

2. P10, line 2: inclusion of claudication patients only may limit generalizability. Although does prove the point about these specific patients.

We consider this to be the main indication for surgery in LSS. Therefore, rather than limiting generalisability, it means that our results will be generalisable to the majority of people who are offered surgery for this condition.

3. P10, inclusions: ‘have not improved’ – is this no improvement at all …? It might help to clarify your meaning here.

There is a pragmatic component to the inclusion criteria. Patients must have not improved sufficiently for both the patient and surgeon to agree that conservative treatment has not been effective and that surgery is indicated.

4. Exclusions: a) There is no psychological assessment included here - is that deliberate? If so, why? b) Might need better definition of CE syndrome – P10, line 35. What is your definition of CE syndrome and how severe would it have to be to be an exclusion?

A baseline the Hospital Anxiety and Depression Scale will be complete by all patients (page 14, under ‘other data collected’). The surgeon will also screen the patient, and must be convinced that surgery has the possibility of benefit taking into account their psychological status as per usual care.

Cauda equina will be routinely diagnosed based on symptoms (i.e. urinary incontinence, numbness and/or tingling in the genital area) and imaging. Given we are only including experienced surgeons, we are confident that CE patients will be effectively identified and excluded.

5. P11, blinding. Is this feasible in all the sites – to keep the blinding working?

We consider participant blinding to be crucial to the success of this trial. We have therefore put in place a number of safety nets to ensure successful blinding at all sites. This includes training of staff at each site in how to conduct a placebo-controlled clinical trial and providing trial information to other clinicians (e.g. general practitioners) involved with the patient care.

6. P13, line 38 Primary outcomes: it is a little concerning that the primary outcome time point will be at 3 months - this is very early to judge the outcome of surgery fairly against conservative (placebo) therapy. There is a possibility of a high number of unsatisfied patients at 3 months wanting to be unblinded with the hope of receiving 'true' surgery. This will then impact on long-term outcome assessment - which is really of more interest to patients than 3-month outcome.

The control group is not conservative therapy, and participants will have already failed conservative therapy. The control group is placebo surgery and given that the response to surgery for this condition is expected to be immediate, we expect recovery to occur within 3 months if surgery is to be considered successful. As noted above, we do not expect a high cross-over and we do agree that longer term outcome is also of interest.

7. P12, line 38: There is no definition of 'poor' clinical outcome. This will need to be more clearly defined to avoid losing a high proportion of patients from blinded assessment at later time points. It would seem to be better to only break blinding for those that are doing very poorly (eg if they have since developed severe CE disease).

Unblinding of participants will occur only in exceptional circumstances, and when knowledge of treatment allocation is essential for further management of the participant. This decision will be made by the Chair of the Steering Committee in consultation with the treating surgeons and clinicians.

8. P15: Would like to see more detail in adverse events – what sort of things, what will you do about them etc.

Standard operating procedures on clinical trial safety reporting will be used. Events will be classified according to their attribution (not related, doubtful, possible, probable, and very likely); and severity (mild; moderate; or severe). Serious adverse events will be any untoward medical occurrence that results in death, is life-threatening, requires hospitalisation or prolongation of existing hospitalisation, persistent or significant disability, or is a medically significant event or reaction. Potential risks and complications of surgical decompression may include cerebrospinal fluid leak, postoperative instability of the operated level, infections, nerve root damage, bleeding and others. While in hospital, adverse events will be considered anything requiring clinical review. Following discharge from hospital, all suspected or confirmed adverse events will be collected and reviewed by an independent medical physician. This information has been clarified in the manuscript (pages 15):

Events will be classified according to their attribution (not related, doubtful, possible, probable, and very likely); and severity (mild; moderate; or severe). Risks and complications of surgical decompression are rare and may include cerebrospinal fluid leak, postoperative instability of the operated level, infection, nerve root damage, and bleeding.

9. P16, line 27: although this will be complicated by the patients' knowledge of treatment group – can argue should not be included.

We are unsure of what the reviewer is referring to.

10. P16, line 37: A bit of a contentious statistical point - there are no true continuous outcomes here! – these are all ordered categories.

The Oswestry Disability Questionnaire, all subscales of the Swiss Spinal Stenosis Questionnaire, the two pain numerical rating scales, and self-reported perceived recovery will be analysed as continuous outcomes.

11. P18, line 24: these are pretty big differences - change in ODI of 15 points in a 50-point scale or between 30% and 60% recovery rate - so negative results may not be enough to definitively persuade

surgeons that surgery is not worthwhile.

The between-group difference of 15 points in the ODI is based on a final score of 100 points (not 50 points – as clarified in the manuscript page 17) and lower than previous published values for minimal clinically important difference in this population12. Likewise, a difference of 30% between groups is based on previous literature12.

References

1. Forsth P, Olafsson G, Carlsson T, et al. A Randomized, Controlled Trial of Fusion Surgery for Lumbar Spinal Stenosis. N Engl J Med 2016;374(15):1413-23.
2. Weinstein JN, Tosteson TD, Lurie JD, et al. Surgical versus nonsurgical therapy for lumbar spinal stenosis. The New England Journal of Medicine 2008;358(8):794-810.
3. Delitto A, Piva SR, Moore CG, et al. Surgery versus nonsurgical treatment of lumbar spinal stenosis: a randomized trial. Ann Intern Med 2015;162(7):465-73. d
4. Pratt RK, Fairbank JC, Virr A. The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. Spine (Phila Pa 1976) 2002;27(1):84-91.
5. Tomkins-Lane CC, Battie MC. Validity and reproducibility of self-report measures of walking capacity in lumbar spinal stenosis. Spine (Phila Pa 1976) 2010;35(23):2097-102.
6. Conway J, Tomkins CC, Haig AJ. Walking assessment in people with lumbar spinal stenosis: capacity, performance, and self-report measures. Spine J 2011;11(9):816-23.
7. Beard DJ, Rees JL, Cook JA, et al. Arthroscopic subacromial decompression for subacromial shoulder pain (CSAW): a multicentre, pragmatic, parallel group, placebo-controlled, three-group, randomised surgical trial. Lancet 2018;391(10118):329-38.
8. Dachun X, Jue L, Liling Z, et al. Sensitivity and specificity of the ankle--brachial index to diagnose peripheral artery disease: a structured review. Vasc Med 2010;15(5):361-9.
9. Cleland JA, Whitman JM, Houser JL, et al. Psychometric properties of selected tests in patients with lumbar spinal stenosis. Spine J 2012;12(10):921-31.
10. Ferreira ML, Herbert RD, Crowther MJ, et al. When is a further clinical trial justified? BMJ 2012;345:e5913.
11. Ferreira ML, Herbert RD, Ferreira PH, 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(3):253-61.
12. Tomkins-Lane CC, Battie MC, Macedo LG. Longitudinal construct validity and responsiveness of measures of walking capacity in individuals with lumbar spinal stenosis. Spine J 2014;14(9):1936-43.

VERSION 2 – REVIEW

REVIEWER	Carlo Ammendolia University of Toronto, Canada
REVIEW RETURNED	18-Oct-2018

GENERAL COMMENTS	Congratulation on your study that address a very important question that impact a large number of people world wide.

REVIEWER	Kazunari Fushimi Department of Orthopaedic Surgery, Gifu University Graduate School of Medicine, Japan.
REVIEW RETURNED	22-Oct-2018

GENERAL COMMENTS	Now, this manuscript is ready for publication. Points requested by reviewers have been appropriately revised. This study addresses an important topic in clinical medicine. All participants and assessors are blinded to treatment allocation. The protocol has been approved by ethics committees. Data analysis is conducted by both randomized controlled trial analysis and intention-to-treat analysis. The authors use an appropriate outcome measures, such as pain analysis (NRS for low back pain and lower limb pain), functional status (ODI), and quality of life (QALYS).