PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	OPAL: A randomised, placebo-controlled trial of opioid analgesia for the reduction of pain severity in people with acute spinal pain. Trial protocol
AUTHORS	Lin, ChungWei Christine; McLachlan, Andrew; Latimer, Jane; Day, Richard; Billot, Laurent; Koes, Bart; Maher, Chris

VERSION 1 - REVIEW

REVIEWER	Danielle van der Windt Keele University, UK
REVIEW RETURNED	31-Mar-2016

GENERAL COMMENTS	This is a well-written protocol for a triple-blind placebo controlled trial on the effectiveness of opioids for spinal pain in primary care. The protocol does not provide a lot of detail on some of the trial procedures (I have highlighted a few areas), and the discussion section could be strengthened. At this stage, the authors may no longer be able to amend the trial protocol itself, but the protocol paper offers a good opportunity justify and discuss some of their decisions and dilemma's in a bit more detail. 1) Title: I would strongly recommend to take out "The first" from the title. This is not relevant or appropriate for the title, and you never know, someone might beat them to it. The title needs to explain study question and design, so do include the term 'randomized', setting of the study (primary care), and is possible the primary outcome (pain severity). 2) Study sample: The trial aims to include patients with 'acute spinal pain'. I am a little bit concerned by the use of the term 'acute' vs 'chronic' pain. Although these terms are still in use, primary care back pain researchers have increasingly acknowledged that spinal pain cannot simply be divided into acute versus chronic pain based on the duration of the pain episode (less or more than 12 weeks).[e.g. Croft 1996; Von Korff & Dunn 2008; Pranski et al. 2010] The prevalence of back pain is very high and increases most strongly during adolescence and early adulthood, with very little change after the age of 30.[Hoy et al. 2012] This means that most adults with spinal pain will have experienced previous episodes of spinal pain, with the number of episodes being a strong predictor of future pain episodes. So if the authors include patients with an acute episode of spinal pain (<12 weeks), they may still recruit a heterogeneous sample including people who have had multiple previous pain episodes, and may have the characteristics of someone with (or at high risk of) persistent recurrent back pain. A one month pain-free period will help to identify a new episode (or exacerbation of mild persistent pain), but will certainly not be able to

	screen out those with recurrent back pain. The authors may not want to change their eligibility criteria, but it would be good to discuss this issue in the protocol paper. People with recurrent spinal pain are likely to have a very different prognosis and potentially a very different response to treatment, in terms of the potential benefits and harms of opioids. 3) Study sample: the study excludes people who have taken a prescription opioid analgesic for the current episode of low back pain and/or neck pain. As above, the sample may still include patients who may have taken opioids for previous episodes of spinal pain, have a different prognosis, a different response to opioids, or may be un-blinded more easily. This needs to be acknowledged in the discussion. 4) Outcome assessment: It would be good to discuss and justify decisions regarding the timing of outcome assessments and choice of outcome measures. For example, there is much uncertainty regarding the effectiveness of opioids in spinal pain in terms of the impact on long-term outcomes, not only pain recurrence, but also function. However, the long-term outcome assessment does not include function or disability. 5) Sample size: The sample size calculation is based on a 1-point difference on the primary outcome of pain severity, which seems realistic and justified. Does this concern pain severity at a specific time-point (6 weeks? 12 weeks?) or based on repeated measures? Does the sample size calculation take into account the repeated measures, and adjustment for baseline value of the outcome measure, proposed in the analysis plan? This may have a large effect on the power (required sample size) of the study.[Vickers 2003] 6) Expected attrition: The sample size calculation is based on max 5% attrition or non-adherence. I realise the Australian team has been highly successful in terms of trial follow-up rates, but this does feel optimistic, and needs justification. 7) Recruitment: I am wondering if it is possible to review recruitment rates in the general practices: can GP records be reviewed to assess what proportion of patients have been invited, and how many have consented to participation out of all those consulting with spinal pain? (This would help to generate the recruitment flowchart). 8) The Steering Committee will consist of the authors of the protocol. Will there be an independent Trial Steering Committee? 9) Blinding: Patients, GPs, and trial personnel will be blind to treatment allocation. For how long will blinding of patients be maintained? Only during the 6-week treatment period or for longer? 10) Data collection: The methods of online data or email collection are only briefly mentioned. Can the authors provide a bit more detail on how this will be organised, and whether or not this has been tested against data collection by telephone? 11) Data collection: All outcomes will be collected by through participant self-report (either by phone via a blinded research assistant or directly by the participant online or by email). I was wondering if prescription data on opioids over the follow-up period of

	12 months could also be retrieved via the GPs medical records. It would be good to also have data on opioid prescription that does not strictly rely on patient self-report. 13) Statistical analysis: In the primary analysis, longitudinal linear models (GEE) will be used to assess the effect of treatment group on pain severity over the 6-week treatment period. Earlier in the protocol the primary outcome is described as pain severity over a period of 12 weeks? GEE assumes missing data are missing completely at random. Other models might perhaps be more appropriate (longitudinal mixed (growth) model?) 14) Sensitivity analyses: These do not include a sensitivity analysis based on treatment adherence. Has this been considered? 15) Based on my previous comment: I am wondering if the authors have considered using previous pain episodes (or previous use of opioids) as co-variate, rather than duration of the current pain episode. 16) Discussion: The discussion is very brief and mainly a repetition of the rationale and objective of the trial. However, this section offers space to discuss some of the dilemma's when designing the trial in terms of the target population, outcome measures, choice of medication, co-interventions, etc. Similarly, potential strengths and weaknesses of the design can be discussed in more detail, which may benefit other researchers designing similar trials.

REVIEWER	Dr Harbinder Kaur Sandhu University of Warwick United Kingdom
REVIEW RETURNED	01-Apr-2016

GENERAL COMMENTS	Overall a well designed and explained protocol. Ethical approval has been obtained and all ethical considerations are reported in the protocol. Clear study rational and methods.

REVIEWER	Laxmaiah Manchikanti, MD. Chief Executive Officer and Chairman of the Board, ASIPP and SIPMS Medical Director, Pain Management Center of Paducah Associate Clinical Professor Anesthesiology and Perioperative Medicine University of Louisville, Kentucky 2831 Lone Oak Road Paducah, KY 42003 USA
REVIEW RETURNED	28-Apr-2016

GENERAL COMMENTS	Well designed control trial- Only concern in that it won't influence medical community similar to Portnoy with 38 patients leading to explosive use of opioids leading to the epidemic. Authors should mention opioid epidemic and deaths. Author should also mention potential for long-term use, even after 1 month. There is no mention of effects on future outcomes in patients treated with opioids.

REVIEWER	Simon Dagenais Spine Research LLC, USA Previous employee, shareholder, and consultant to Palladian Health, a private health management company involved in musculoskeletal programs for low back pain, neck pain, and other conditions
REVIEW RETURNED	09-May-2016

GENERAL COMMENTS	Abstract: line 14: consider adding "strong" before "efficacy data" line 34: consider defining "guideline care" line 36: consider listing effectiveness outcomes Manuscript page 5/line 52: consider adding "for acute low back pain and neck pain" after "to support their use" page 6/line 30: consider deleting "quality" before "use" page 6/line 37: consider explaining the use of a placebo control rather than an active control (APAP + NSAIDs) in terms of clinical equipoise and feasibility of recruitment. If this trial is intended to be practical and reflect primary care decisions, the use of placebo is inappropriate. page 7/line 24: having 200 study sites introduces a potentially large source of variation in the interventions (e.g. education). How will this be managed and minimized? page 7/line 41: consider limiting enrollment to mechanical/uncomplicated acute low back or neck pain rather than a broader pain with or without radiculopathy page 7/line 50: define how a "pain-free period" will be assessed and determined page 8/line 14: explain how the exclusion criteria will be assessed (e.g. mandatory diagnostic imaging/laboratory testing for all participants?) page 8/line 30: are any opioids available over the counter in Australia? page 9/line 39: "cannot stay away" should likely be "cannot stay awake" page 10/19: self-reported use of opioids will likely not be valid to detect misuse/abuse. Consider urine testing to detect opioid use after 6 weeks. page 10/line 30: Describe exactly how guideline-based care will be delivered, as this is no small task. Will advice be standardized using booklets? Videos? Telephone support? Who will offer spinal manipulative therapy? How often? For how long? page 11/line 30: why not use the Oswestry Disability Index, which is the low back equivalent of the Neck Disability Index, rather than the Roland-Morris? Also, why not measure physical functioning at 12 weeks? page 11/line 39: how will adherence to a pain diary be enforced? page 12/line 25: why not measure NDI/ODI/SF-12 at 3, 6 and 12 months? page 17/line 12: are there utility weights for an Australian population to convert SF-12 scores into QALYs? page 17/line 39: self-reported absence from work due to back or neck pain may be an outcome with poor validity

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Danielle van der Windt
Institution and Country: Keele University, UK Competing Interests: None declared

Reviewer 1's comment 1
1) Title: I would strongly recommend to take out "The first" from the title… The title needs to explain study question and design, so do include the term 'randomized', setting of the study (primary care), and is possible the primary outcome (pain severity).

Authors' response 1
We have amended the title following the suggestions from the reviewer and the SPIRIT 2013 Statement (Chan et al 2013, BMJ;346:e7586). We have removed "The first" from the title and added the term "randomised", as well as the primary outcome.

Reviewer 1's comment 2
2) Study sample: The trial aims to include patients with 'acute spinal pain'. I am a little bit concerned by the use of the term 'acute' vs 'chronic' pain. Although these terms are still in use, primary care back pain researchers have increasingly acknowledged that spinal pain cannot simply be divided into acute versus chronic pain … most adults with spinal pain will have experienced previous episodes of spinal pain, with the number of episodes being a strong predictor of future pain episodes. So if the authors include patients with an acute episode of spinal pain (<12 weeks), they may still recruit a heterogeneous sample including people who have had multiple previous pain episodes, and may have the characteristics of someone with (or at high risk of) persistent recurrent back pain…
The authors may not want to change their eligibility criteria, but it would be good to discuss this issue in the protocol paper. People with recurrent spinal pain are likely to have a very different prognosis and potentially a very different response to treatment, in terms of the potential benefits and harms of opioids.

Authors' response 2
We agree that there are ways of grouping people with non-specific spinal pain other than by symptom duration. We have designed the OPAL trial in an acute population for 2 main reasons. Firstly, there are already trials in the chronic population but none in the acute population. Secondly, spinal pain guidelines currently make recommendations based on whether the condition is acute or chronic (and the Cochrane Back and Neck Group organise their reviews in the same way). By including people with acute spinal pain the OPAL trial can address an evidence gap as well as directly inform clinical practice and guidelines.

We also agree that our inclusion criteria will capture a heterogeneous population. Currently there are no randomised, placebo-controlled trials on opioid analgesics for acute spinal pain. So the first step is to conduct a robust trial to establish efficacy and safety of opioid analgesics to the wide range of people with spinal pain that present to primary care. If the OPAL trial shows that opioid analgesics are efficacious and well-tolerated in people with acute spinal pain, then future research can investigate treatment responsiveness according to factors such as patient prognostic factors and medication dosage. We have added the following to the discussion.

> *"The inclusion and exclusion criteria have been kept as broad as possible so the results can be applicable to the heterogeneous population that present to primary care with acute, non-specific spinal pain. OPAL is specifically designed to establish the effects of opioid analgesics in acute spinal pain. If the results show that opioid analgesics are indeed efficacious and well-tolerated, future research can further investigate if factors such as patient prognostic characteristics, treatment expectation, medication dosage and the type of opioid analgesics used will influence treatment responsiveness."*

Reviewer 1's comment 3
3) Study sample: the study excludes people who have taken a prescription opioid analgesic for the current episode of low back pain and/or neck pain. As above, the sample may still include patients who may have taken opioids for previous episodes of spinal pain, have a different prognosis, a

different response to opioids, or may be un-blinded more easily. This needs to be acknowledged in the discussion.

Authors' response 3
The value of a randomised controlled trial is that we can use randomisation to balance treatment groups with respect to known and unknown prognostic factors.

In our previous review of opioids for spinal pain we did not find evidence that there was a problem with unblinding in the placebo controlled trials (Shaheed et al. JAMA Intern Med. 2016 May 23. doi: 10.1001/jamainternmed.2016.1251). We will assess the success of blinding and if there are differences between groups will consider that in our interpretation of the results.

We acknowledge the possibility of differential response to treatment (ie treatment effect modification) but to explore this reliably we would need a far larger trial and we do not the have the budget for this.

We have included a comment on these issues in the revised manuscript.

Reviewer 1's comment 4
4) Outcome assessment: It would be good to discuss and justify decisions regarding the timing of outcome assessments and choice of outcome measures. For example, there is much uncertainty regarding the effectiveness of opioids in spinal pain in terms of the impact on long-term outcomes, not only pain recurrence, but also function. However, the long-term outcome assessment does not include function or disability.

Authors' response 4
We agree that long term outcomes would be important, but only if we were testing long term use of the medicine. Our trial tests brief use of this medicine; up to a maximum of 6 weeks. We have stated that the outcomes and outcome measures are those recommended by consensus of international experts in the IMMPACT recommendations (please see page 11), and have added the following paragraph to the discussion:

> *"We have chosen outcomes that are patient-oriented, clinically relevant and recommended by international consensus.[28] Because the trial intervention will be implemented over a short period of time for an acute condition where patients in general make an initial, rapid improvement,[36 40] we will collect outcomes frequently (every 2 weeks) during the 6-week treatment period to assess treatment effects. We will also collect outcomes at 12 weeks to assess whether treatment effects, if any, can be sustained in the medium term. While we do not expect that a short term treatment will influence long term clinical outcomes, we will collect our primary outcome at 6 and 12 months, as well as long term opioid use and misuse data, to inform our results."*

We have not planned to assess many long term clinical outcomes (including function or disability), other than the primary outcome and use and misuse data, because we do not expect that a short term treatment will have long term clinical implications.

Reviewer 1's comment 5
5) Sample size: The sample size calculation is based on a 1-point difference on the primary outcome of pain severity, which seems realistic and justified. Does this concern pain severity at a specific time-point (6 weeks? 12 weeks?) or based on repeated measures? Does the sample size calculation take into account the repeated measures, and adjustment for baseline value of the outcome measure, proposed in the analysis plan? This may have a large effect on the power (required sample size) of the study.[Vickers 2003]

Authors' response 5
Our intention is for the primary analysis to include pain severity at all time points up to Week 6 using a repeated-measure linear mixed model; however, the sample size calculation was based on a single time-point (6 weeks) and does not account for the baseline measure. This was done on purpose to be conservative as well as to acknowledge the uncertainty around a) the correlation between repeated measures and b) the correlation between the baseline value and the value during follow-up.

We have now specified in the manuscript that the sample size calculation is based on effect at 6 weeks.

Reviewer 1's comment 6
6) Expected attrition: The sample size calculation is based on max 5% attrition or non-adherence. I realise the Australian team has been highly successful in terms of trial follow-up rates, but this does feel optimistic, and needs justification.

Authors' response 6
We have in fact allowed 5% dropout as well as 10% non-compliance (please see section on 'Sample size' on page 13 of the manuscript). The 5% dropout is based on our previous experience, for example in the PACE study we included 99.5% of the sample in the analysis of the primary outcome (Williams et al 2014, Lancet 384,1586-1596). We have now provided a brief justification.

"A sample size of 173 per group …, assuming a standard deviation of 2.5,[22] power of 90% and alpha of 5%, and allowing for 5% dropout (based on our previous study)[22] …"

Reviewer 1's comment 7
7) Recruitment: I am wondering if it is possible to review recruitment rates in the general practices: can GP records be reviewed to assess what proportion of patients have been invited, and how many have consented to participation out of all those consulting with spinal pain? (This would help to generate the recruitment flowchart).

Authors' response 7
We have asked GPs to screen all potentially eligible participants to assess what proportion of participants consent to participating in the trial. But we have not planned to review GP records due to the resources that this will take. Due to privacy and confidentiality reasons, trial staff will not be able to access GP records so obtaining such information will require staffing and time from the GP practice and likely impact on a GP's willingness to participate in the trial. Unfortunately in Australia we do not have the computerised system that primary care researchers in the UK use.

Reviewer 1's comment 8
8) The Steering Committee will consist of the authors of the protocol. Will there be an independent Trial Steering Committee?

Authors' response 8
The Steering Committee will make the final decisions and assume responsibility for the trial so will need to be investigators on the trial. However we have convened an independent data and safety monitoring board (please see section on 'Data monitoring' on page 18 of the manuscript).

Reviewer 1's comment 9
9) Blinding: Patients, GPs, and trial personnel will be blind to treatment allocation. For how long will blinding of patients be maintained? Only during the 6-week treatment period or for longer?

Authors' response 9
Blinding to patients, GPs and trial personnel will be maintained for the entire duration of the trial to maintain integrity. We have added the following sentence to the section on 'Blinding' (page 14):

"Blinding will be maintained for the entire duration of the trial until all data have been collected and data analysis and interpretation have been completed and agreed upon."

Reviewer 1's comment 10
10) Data collection: The methods of online data or email collection are only briefly mentioned. Can the authors provide a bit more detail on how this will be organised, and whether or not this has been tested against data collection by telephone?

Authors' response 10
We have added more information to the section on data collection:

'*For participants completing data assessment online, an email reminder will be sent prior to the due date, with a unique link to complete the data online. For participants completing data assessment by email, a blank assessment form will be provided which can be emailed back to the trial team upon completion.*'

We have not planned to formally test the data collected online or by email against those collected by phone.

Reviewer 1's comment 11
11) Data collection: All outcomes will be collected by through participant self-report ... I was wondering if prescription data on opioids over the follow-up period of 12 months could also be retrieved via the GPs medical records. It would be good to also have data on opioid prescription that does not strictly rely on patient self-report.

Authors' response 11
We have indeed intended to collection prescription data on study medication during the treatment period, and apologise for omitting such information. This information is now added to the sections on 'Outcomes' (addition underlined):

> '*Adherence to trial medication – measured by participants' self-report of daily trial medication intake recorded in a diary and by counting the returned medications. <u>This will be compared against prescription data of trial medication reported by the general practitioner at each participant visit during the6-week treatment period</u>.*'

And 'Data collection methods':

> *'Additionally, at each participant visit general practitioners will be asked to complete data on trial medication prescription and adverse events (if any, excluding the initial visit).*'

Collecting ongoing opioid use following the 6-week treatment period from a more objective source such as the GP's medical records will strengthen the trial but we have not planned for this within the current design, due to resource implications and because opioid misuse is rare (Minozzi et al 2013, Addiction;108:688-98; Noble et al 2010, Cochrane Database Syst Rev;Issue 1:CD006605.) especially when using opioids for a limited time for acute pain. We acknowledge that this is a potential weakness to the trial and have added this to the discussion.

Reviewer 1's comment 13 (no comment 12)
13) Statistical analysis: In the primary analysis, longitudinal linear models (GEE) will be used to assess the effect of treatment group on pain severity over the 6-week treatment period. Earlier in the protocol the primary outcome is described as pain severity over a period of 12 weeks?
GEE assumes missing data are missing completely at random. Other models might perhaps be more appropriate (longitudinal mixed (growth) model?)

Authors' response 13
The primary aim is to investigate pain severity over 6, not 12, weeks (see page 6). Pain severity at 12 weeks is part of the secondary analysis.

We have changed the analysis to repeated-measure linear mixed models and have updated the manuscript accordingly.

Reviewer 1's comment 14
14) Sensitivity analyses: These do not include a sensitivity analysis based on treatment adherence. Has this been considered?

Authors' response 14
We have not planned to conduct a per-protocol or a sensitivity analysis based on treatment adherence. If appropriate we may conduct a post-hoc analysis (e.g. a Complier-average Causal Effect (CACE) analysis), but this will be very exploratory and limited by the small numbers.

Reviewer 1's comment 15
15) Based on my previous comment: I am wondering if the authors have considered using previous pain episodes (or previous use of opioids) as co-variate, rather than duration of the current pain episode.

Authors' response 15
Our primary analysis will only include baseline pain severity as a co-variate. Pain duration and site of pain will be additional covariates in a sensitivity analysis to support the main analysis (please see

page 16 of the manuscript and our response to Reviewer 1's comment 2). We have not planned to include previous pain episodes or previous use of opioids as covariates.

Reviewer 1's comment 16
16) Discussion: The discussion is very brief and mainly a repetition of the rationale and objective of the trial. However, this section offers space to discuss some of the dilemma's when designing the trial in terms of the target population, outcome measures, choice of medication, co-interventions, etc. Similarly, potential strengths and weaknesses of the design can be discussed in more detail, which may benefit other researchers designing similar trials.

Authors' response 16
Thank you for the suggestion. The discussion has been expanded to cover the issues mentioned by the reviewer.

Reviewer: 2
Reviewer Name: Dr Harbinder Kaur Sandhu
Institution and Country: University of Warwick, United Kingdom Competing Interests: None

Overall a well designed and explained protocol. Ethical approval has been obtained and all ethical considerations are reported in the protocol. Clear study rational and methods.

Authors' response
None required.

Reviewer: 3
Reviewer Name: Laxmaiah Manchikanti, MD.
Institution and Country: Chief Executive Officer and Chairman of the Board, ASIPP and SIPMS; Medical Director, Pain Management Center of Paducah; Associate Clinical Professor Anesthesiology and Perioperative Medicine University of Louisville, Kentucky, USA Competing Interests: None declared

Well designed control trial- Only concern in that it won't influence medical community similar to Portnoy with 38 patients leading to explosive use of opioids leading to the epidemic. Authors should mention opioid epidemic and deaths. Author should also mention potential for long-term use, even after 1 month. There is no mention of effects on future outcomes in patients treated with opioids.

Authors' response
As mentioned in the authors' response to Reviewer 1's comment 2, we have designed the trial to address an evidence gap as well as directly inform clinical practice and guidelines, so we do expect that the results will be influential and contribute to evidence-based practice. Our previous trials evaluating medicines have been highly influential (e.g. Hancock et al (2007). Lancet 370,1638-1643; Williams et al (2014). Lancet 384,1586-1596).

We have mentioned the widespread use of opioids and potential for misuse (pages 5 and 6). We have now added information about overdose and deaths to paragraph 4 of the introduction.

> *"Australian data suggest that hospital separations for opioid poisoning doubled between 2005 to 2006 and 2006 to 2007,[15] and there has been a multi-fold increase in drug overdose deaths in the United States in the last 15 years, primarily driven by opioid analgesics.[16] Alarmingly, since 2003 more overdose deaths have involved opioid analgesics than heroin and cocaine combined.[17]"*

We will collect long term outcomes on the primary outcome of pain severity, and opioid use and misuse at 1 year.

Reviewer: 4
Reviewer Name: Simon Dagenais

Institution and Country: Spine Research LLC, USA Competing Interests: Previous employee, shareholder, and consultant to Palladian Health, a private health management company involved in musculoskeletal programs for low back pain, neck pain, and other conditions

Reviewer 4 comment 1
Abstract:
line 14: consider adding "strong" before "efficacy data"
line 34: consider defining "guideline care"
line 36: consider listing effectiveness outcomes

Authors' response 1
We have added "strong" before "efficacy data". We have re-termed "guideline care" as "guideline-based care" and given an example (*"e.g. reassurance and advice of staying active"*) to make it clearer. The primary outcome was already listed after the sentence ending "effectiveness outcomes". We have now also given examples of secondary outcomes. The relevant section reads (addition underlined):

> *"Participants will be followed for 3 months for effectiveness outcomes. The primary outcome will be pain severity. Secondary outcomes will include physical functioning and time to recovery."*

Reviewer 4 comment 2
Manuscript
page 5/line 52: consider adding "for acute low back pain and neck pain" after "to support their use"
page 6/line 30: consider deleting "quality" before "use"

Authors' response 2
We have added "for acute low back pain and neck pain" and deleted "quality" before "use".

Reviewer 4 comment 3
page 6/line 37: consider explaining the use of a placebo control rather than an active control (APAP + NSAIDs) in terms of clinical equipoise and feasibility of recruitment. If this trial is intended to be practical and reflect primary care decisions, the use of placebo is inappropriate.

Authors' response 3
We have added a sentence to the section on 'Design':

> *"We have chosen a randomised, placebo-controlled design in order to provide high-level evidence on the efficacy or safety of opioid analgesics in those with acute spinal pain."*

Our trial is not intended to **reflect** primary care practice because of course placebos are not dispensed in primary care. The value of our trial is that it will rigorously evaluate the efficacy of this medicine and so will **inform** primary care practice and the use of a placebo is key to reducing bias.

We urgently need a placebo-controlled trial of opioid analgesics in acute spinal pain to inform the debate of whether using this medication is justified in this population.

Reviewer 4 comment 4
page 7/line 24: having 200 study sites introduces a potentially large source of variation in the interventions (e.g. education). How will this be managed and minimized?

Authors' response 4

> *"General practitioners will be trained on the trial protocol and receive regular monitoring visits from the trial team to ensure adherence to the trial intervention."* (page 9)

Reviewer 4 comment 5
page 7/line 41: consider limiting enrolment to mechanical/uncomplicated acute low back or neck pain rather than a broader pain with or without radiculopathy page 7/line 50: define how a "pain-free period" will be assessed and determined page 8/line 14: explain how the exclusion criteria will be assessed (e.g. mandatory diagnostic imaging/laboratory testing for all participants?) page 8/line 30: are any opioids available over the counter in Australia?

Authors' response 5
We will keep the inclusion criteria to acute low back pain or neck pain with or without radiation to the leg or arm, but will collect information on whether participants have pain radiation or not. The pain-free period and the exclusion criteria will be assessed via a history and physical examination. This is now stated in the protocol in the section 'Eligibility criteria'.

Currently some combination weak opioids are available over counter in Australia, e.g. codeine 8mg plus paracetamol 500mg (trade name panadeine).

Reviewer 4 comment 6
page 9/line 39: "cannot stay away" should likely be "cannot stay awake"
page 10/19: self-reported use of opioids will likely not be valid to detect misuse/abuse. Consider urine testing to detect opioid use after 6 weeks.

Authors' response 6
The spelling mistake has been corrected to "awake". We acknowledge that self-reported use of opioids can be subject to reporting bias. However balancing trial resources, the recruitment of an acute population with short term use of opioids, and exclusion of those with a history or at high risk of addiction, drug abuse or alcoholism, we have chosen self-report to monitor misuse/abuse. We have now addressed this limitation in the discussion.

Reviewer 4 comment 7
page 10/line 30: Describe exactly how guideline-based care will be delivered, as this is no small task. Will advice be standardized using booklets? Videos? Telephone support? Who will offer spinal manipulative therapy? How often? For how long?

Authors' response 7
We have expanded the paragraph on guideline-based care to make it clearer (additions underlined).

> *"In addition to the trial medication, general practitioners will deliver guideline-based care to all participants.[6] This includes advice (patient reassurance, staying active and avoiding bed rest) delivered by the general practitioners and, if required, other guideline-recommended treatments delivered either by the general practitioners or by referral to other health professionals. For example, spinal manipulative therapy may be offered and simple analgesics or adjuvants may be provided in addition to the trial medication in accordance to the WHO analgesic ladder.[6 27] However, no concomitant opioid analgesics should be used. We have not protocolised guideline-based care, other than that all participants will receive advice, to allow the care to be tailored to participants."*

Reviewer 4 comment 8
page 11/line 30: why not use the Oswestry Disability Index, which is the low back equivalent of the Neck Disability Index, rather than the Roland-Morris? Also, why not measure physical functioning at 12 weeks?
page 11/line 39: how will adherence to a pain diary be enforced?
page 12/line 25: why not measure NDI/ODI/SF-12 at 3, 6 and 12 months?
page 17/line 12: are there utility weights for an Australian population to convert SF-12 scores into QALYs?
page 17/line 39: self-reported absence from work due to back or neck pain may be an outcome with poor validity

Authors' response 8
The Oswestry disability index and Roland Morris questionnaire are both widely used and validated questionnaires. We have chosen the Roland Morris questionnaire because it covers more functional tasks. The Roland Morris questionnaire and the Neck Disability Index will be used as condition-specific physical functioning measures and will only be applied to participants with low back pain and neck pain respectively. We will also measure physical functioning using a generic scale (The Pain Interference Score of the Brief Pain Inventory), which can be administered to all participants and will be assessed at 2, 4, 6 and 12 weeks. Therefore, to reduce participant burden, we will only administer the condition-specific physical functioning measures at the main time points of baseline and 6 weeks (end of treatment). We are not measuring the Roland Morris questionnaire, the Neck Disability Index or the SF-12 at 6 and 12 months as we consider short term use of an analgesic unlikely to influence

long term clinical outcomes (please see authors' response to Reviewer 1's comment 4). Please note that SF-12 will be collected at 12 weeks/3 months.

Adherence to a pain diary will be encouraged by research assistants collecting data from participants.

There are utility weights for an Australian population for SF-12 (Norman et al (2014). Valuing SF-6D health states using a discrete choice experiment. Medical Decision Making 34,773-786.).

Self-reported absence of work is subject to the same bias as other self-reported data.

VERSION 2 – REVIEW

REVIEWER	Simon Dagenais Spine Research, LLC Previous employee, shareholder, and consultant to Palladian Health, a private health management company involved in musculoskeletal programs for low back pain, neck pain, and other conditions
REVIEW RETURNED	16-Jun-2016

GENERAL COMMENTS	The authors have addressed most of the comments raised by reviewers and improved this manuscript.