Dear Drs. Godlee, Burch and Members of the Editorial Board,

On behalf of all my co-authors, we are grateful for the opportunity to revise our manuscript for BMJ, and have done our best to address every one of the comments listed below. We continue to believe that this will be a groundbreaking article (there are many guidelines that come out on sciatica, ESI, gabapentin and other adjuvants, but none have included any double-blind comparative-efficacy studies) and that BMJ is the best venue for this article. We hope that you find are changes sufficient (we would of course be willing to make any additional changes you see fit) and look forward to future correspondence.

Sincerely,

Steven Cohen and Co-Authors

* We require trials to be prospectively registered in accordance with ICMJE guidelines. This trial was reported as registered in the same month that treatment began. Please provide exact dates of the start of patient recruitment and trial registration.

We have noted the date the trial was registered and when recruitment began. These changes have been noted on the cover page and page 6.

*Editors felt that inclusion and exclusion criteria were not clear. The mechanisms for neurogenic claudication and lumbosacral radicular pain are different, so allowing either type of patient may have introduced heterogeneity. How was this handled? What imaging criteria were required?

The authors went over previous randomized ESI and gabapentin studies. The large majority included patients with both spinal stenosis and herniated disc, without making a distinction. Sometimes patients have herniated discs causing spinal stenosis. We required that patients without spinal stenosis have a herniated disc, and that those with neurogenic claudication have spinal stenosis. The actual course of spinal stenosis is not *that* different from HNP in that only a minority of patients clinically progress, though unlike HNP the pathology does not 'regress'. The studies examining whether one responds better than another have yielded conflicting findings. Only a relatively small proportion of patients in our study had spinal stenosis, and a subgroup analysis (noted on page 12) did not come close to finding any differences. The required imaging had to explain the symptoms (e.g. a herniated disc for radicular pain or spinal stenosis for neurogenic claudication). This clarification was made on page 6.

*There was significant confusion about inconsistencies in "back pain" vs "leg pain" in the title, introduction, inclusion criteria, and outcomes. Which of the two was the target of this study? Why are these two different terms used throughout the paper?

In the title, abstract, introduction and methods, we have changed "neuropathic back pain" to "sciatica (a more general term). The term "neuropathic back pain is very confusing. The First Author recently did a symposium with Dr. Rolf Baron at World Institute of Pain last year (he is the author of several of the referenced papers and the Senior Author on the painDETECT validation study) who believes there may be

cases of 'axial-only' back pain that are neuropathic in nature, and of course not all leg pain is neuropathic (can be mechanical pain referred from the facet joints, sacroiliac joint, or discs). However, most instruments (painDETECT, s-LANSS) categorize radiating pain (i.e. into the leg) as likely neuropathic. Because patients cannot accurately distinguish between mechanical and neuropathic pain scores, leg (indicative of neuropathic pain) and back pain scores are always used instead.

It should be known that the term "sciatica" is considered by some organizations to be an anachronism that should be abandoned (https://www.iasp-pain.org/files/Content/ContentFolders/Publications2/ClassificationofChronicPain/PART_I-C.pdf), though it continues to be used. One of our co-authors felt we should not perpetuate the term's use, though the others did not agree. One cannot ever be sure that leg pain is truly "radicular" (degenerated discs, facet and SI joints often radiate into the leg in a sclerotomal distribution but that is mechanical pain). We could have given patients painDETECT each time they came in to increase the likelihood that the leg pain was radicular, but some studies now question the accuracy of these instruments. In our defense, we went to great lengths to confirm that these patients truly had radicular pain by requiring physical exam signs (straight leg raising test) or symptoms consistent with radicular plain + MRI corroboration (elaborated on page 6). We would be willing to change it to radicular pain if the Editors feel that is best.

*There is no rationale provided for the target outcome of 1 point difference on the pain scale. Is this clinically meaningful?

We apologize here. As you can see from both protocols we sent you (and our sister article, the comparative-effectiveness study evaluating cervical ESI that was published in ANESTHESIOLOGY (BMJ had been concerned about the open-label aspect), a 2-point decrease in pain coupled with a positive GPE was always considered to be a positive outcome in this study (this is now noted). A clinically meaningful reduction in pain has been shown to be about 30%, or a 2-point decrease in pain on a 0-10 scale (see Farrar et al. Pain 2001; Dworkin et al. J Pain 2008- IMMPACT recommendations on clinically important outcomes in clinical pain trials, which included members of academia (surgery, psych, neurology), government (FDA) and industry). But the authors of the IMMPACT guidelines note (verbatim), "… in evaluating a new analgesic, if a 2-point decrease on a 0 to 10 NRS of pain intensity is considered a clinically important improvement for an individual, it *should not be inferred that a 2-point difference in pain reduction between the analgesic and placebo must occur before the treatment benefit can be considered clinically important*." We considered a 1-point difference in pain scores as the basis for our power analysis. Most drugs that are approved in the U.S. by the FDA "beat" placebo in 2 trials by between 10% and 15%. In our study, we expected both groups to improve, so a 1-point difference in a lower pain score would be more than a 15% difference. To provide some context, in the Friedly et al. ESI article in N Engl J Med 2014 powered their study for a 0.8 difference between groups (epidural steroids and epidural local anesthetic). The Iverson et al. ESI study in BMJ in 2011 used a 10-point difference in a 0-100 mm scale (same as a 1-point difference in an 11-point scale) between either treatment group and the sham group for their power analysis. In the 1998 JAMA study by Rowbotham et al. that compared gabapentin to placebo in PHN, they used a 1.5 point difference between groups.

*Editors were concerned about the lack of a complete sham/placebo group, rendering the assay sensitivity in doubt.

We agree with the editors, and have added the following to the "limitations" section: "… and the lack of a true placebo group, which renders the assay sensitivity questionable. Without a true placebo group, one cannot assess the true efficacy of the 2 treatments." The reason for we did this was that we felt that there were over 50 placebo-controlled trials comparing ESI or gabapentin to placebo for sciatica, and that (per HHS recommendations) there was a strong need for "comparative-effectiveness" research.

*Editors did not agree with the conclusion that treatment decisions should be based on patient considerations as this was not evaluated in this study. This study only provides evidence that there is no meaningful difference between the two treatment options.

We have omitted that sentence and replaced it with the following sentence on page 17: "The similar outcomes between treatment groups on most measures suggest that a trial with neuropathic medications is a reasonable first-line treatment option before attempting an ESI."

*We felt that the dichotomy of ESI vs. gabapentin may not reflect real life where patients often try a combination of interventions.

We discussed this point on pages 15-16 of the discussion: "In a recent open-label 3-arm comparative-effectiveness study pitting a series of ESI against conservative therapy consisting of pharmacotherapy (gabapentin and/or nortriptyline) and physical therapy and the combination of the two, it was found that combination treatment with ESI plus medication and physical therapy provided superior benefit to stand-alone treatment on some outcome measures. In some respects, the open-label format and inclusion of a multimodal treatment approach may better reflect "real-life" circumstances, though they preclude the evaluation of efficacy.")

*We do not agree with the reviewer who wanted more extrapolation from statistically non-significant data. We consider results that are statistically non-significant to be negative outcomes.

We have left any changes to that effect out.

* We had a number of concerns about the statistical elements and reporting of this trial, all of which are mentioned in the statistical review at the end of this letter (Review 4). We would need to see all of these concerns addressed in entirety prior to reconsidering a revised version of this paper.

We have addressed all of those concerns.

IMPORTANT
When you revise and return your manuscript, please take note of all the following points about revising your article. Even if an item, such as a competing interests statement, was present and correct in the original draft of your paper, please check that it has not slipped out during revision.

a. In your response to the reviewers and committee please provide, point by point, your replies to the comments made by the reviewers and the editors, and please explain how you have dealt with them in the paper. It may not be possible to respond in detail to all these points in the paper itself, so please do so in the box provided

b. If your article is accepted it will then be edited, proofed, and - after your approval - published on bmj.com with open access. This open access Online First article will not be a pre-print. It will represent the full, citable, publication of that article. The citation will be year, volume, elocator (a unique identifier for that article): eg BMJ 2008;337:a145 — and this is what will appear immediately in Medline, PubMed, and other bibliographical indexes. We will give this citation in print and online, and you will need to use it when you cite your article.

c. Please write an abridged version of the article for the print and iPad BMJ using the appropriate BMJ pico template for your study's design. Please be reassured that it doesn't take long to complete this. When your BMJ pico is ready please email it to papersadmin@bmjgroup.com. The templates for you to download are at
http://resources.bmj.com/bmj/authors/bmj-pico

This has been done and will be e-mailed this week.

d. Please include these items in the revised manuscript to comply with BMJ style:

Title: this should include the study design eg "systematic review and meta-analysis"

Abstract
structured abstract including key summary statistics, as explained below (also
see http://resources.bmj.com/bmj/authors/types-of-article/research)
for every clinical trial - and for any other registered study - the study registration number and name of register – in the last line of the structured abstract.

Introduction
this should cover no more than three paragraphs, focusing on the research question and your reasons for asking it now

Methods:
for an intervention study the manuscript should include enough information about the intervention(s) and comparator(s) (even if this was usual care) for reviewers and readers to understand fully what happened in the study. To enable readers to replicate your work or implement the interventions in their own practice please also provide (uploaded as one or more supplemental files, including video and audio files where appropriate) any relevant detailed descriptions and materials. Alternatively, please provide in the manuscript urls to openly accessible websites where these materials can be found
Results
please report statistical aspects of the study in line with the Statistical Analyses and Methods in the Published Literature (SAMPL) guidelines http://www.equator-network.org/reporting-guidelines/sampl/

summary statistics to clarify your message. Please include in the results section of your structured abstract (and, of course, in the article's results section) the following terms, as appropriate:

For a clinical trial:
- Absolute event rates among experimental and control groups

- RRR (relative risk reduction)
- NNT or NNH (number needed to treat or harm) and its 95% confidence interval (or, if the trial is of a public health intervention, number helped per 1000 or 100,000).

NNT information has been included.

For a cohort study:
- Absolute event rates over time (eg 10 years) among exposed and non-exposed groups
- RRR (relative risk reduction)

For a case control study:
- OR (odds ratio) for strength of association between exposure and outcome

For a study of a diagnostic test:
- Sensitivity and specificity
- PPV and NPV (positive and negative predictive values)

one or more references for the statistical package(s) used to analyse the data, eg RevMan for a systematic review. There is no need to provide a formal reference for a very widely used package that will be very familiar to general readers eg STATA, but please say in the text which version you used
for articles that include explicit statements of the quality of evidence and strength of recommendations, we prefer reporting using the GRADE system
Discussion
please write the discussion section of your paper in a structured way, to minimise the risk of careful explanation giving way to polemic.Please follow this structure:
statement of principal findings of the study
strengths and weaknesses of the study
strengths and weaknesses in relation to other studies, discussing important differences in results and what your study adds. Whenever possible please discuss your study in the light of relevant systematic reviews and meta-analyses (eg Cochrane reviews)
meaning of the study: possible explanations and implications for clinicians and policymakers and other researchers; how your study could promote better decisions
unanswered questions and future research

Footnotes and statements

What this paper adds/what is already known box (as described
at http://resources.bmj.com/bmj/authors/types-of-article/research)

ID of ethics committee approval and name of the ethics committee/IRB; or a statement that approval was not required (see http://resources.bmj.com/bmj/authors/editorial-policies/guidelines) and a statement that participants gave informed consent before taking part

a statement that any identifiable patients have provided their signed consent to publication. Please submit, as a supplemental file, the signed BMJ patient consent form giving consent to publication in The BMJ of any information about identifiable individual patients. Publication of any personal information about a patient in The BMJ, for example in a case report or clinical photograph, will normally require the signed consent of the patient.

competing interests statement (see http://resources.bmj.com/bmj/authors/editorial-policies/competing-interests)

contributorship statement+ guarantor (see http://resources.bmj.com/bmj/authors/article-submission/authorship-contributorship)

transparency statement: a statement that the lead author (the manuscript's guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies are disclosed.

The transparency statement and "guarantor" have been added on the cover page.

copyright statement/ licence for publication (see http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyright-open-access-and-permission-reuse)

signed patient consent form(s), if the article gives enough personal information about any patient(s): this sometimes occurs even in research papers - for example in a table giving demographic and clinical information about a small subgroup in a trial or observational study, or in quotes/tables in a qualitative study - (see http://resources.bmj.com/bmj/authors/editorial-policies/copy_of_patient-confidentiality)

a data sharing statement declaring what further information and data you are willing to make available, over and above the results reported in the paper. Suggested wording: "Data sharing: technical appendix, statistical code, and dataset [state whether any patient level data have been anonymised] are available at this repository or website OR from the corresponding author at ". If there are no such further data available, please use this wording: "Data sharing: no additional data available". For papers reporting the main results of trials of drugs or devices we require that the authors state, at a minimum, that the relevant anonymised patient level data are available on reasonable request from the authors
The BMJ has partnered with the Dryad Digital Repository datadryad.org to make open deposition easy and to allow direct linkage by doi from the dataset to The BMJ article and back - we encourage authors to use this option
funding statement (see http://resources.bmj.com/bmj/authors/article-submission/article-requirements)
statement of the independence of researchers from funders
(see http://resources.bmj.com/bmj/authors/article-submission/article-requirements)
for studies funded or sponsored by industry (see http://resources.bmj.com/bmj/authors/article-submission/article-requirements)
a statement describing the role of the study sponsor(s), if any, in study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication
assurance, in the cover letter, that a clinical trial funded by a pharmaceutical or other commercial company follows the guidelines on good publication practice
(seehttp://resources.bmj.com/bmj/authors/article-submission/article-requirements)
inclusion in the list of contributors the name(s) any professional medical writer(s), specifying in the formal funding statement for the article who paid the writer. Writers and authors must have access to relevant data while writing articles.

Patient centered research

for studies that are relevant to patients we expect authors to report in their articles the extent of their study's patient-centredness, as highlighted by these questions:
did you involve patients/service users/carers/lay people in the design of this study? Please state whether you did, and give details (Methods section)
was the development and/or selection of outcome measures informed by patients' priorities and experiences? Please give details (Methods section)
were patients/service users/carers/lay people involved in developing plans for participant recruitment and study conduct? If so, please specify how (Methods section)
have you planned to disseminate the results of the study to participants? If so how will this be done? (Describe in brief footnote)
are patients thanked in the contributorship statement or acknowledgements?
for articles reporting randomised controlled trials: did you assess the burden of the intervention on patients' quality of life and health? If so, what evaluation method did you use, and what did you find? (Methods and Results sections)

We have added an acknowledgment section which includes thanking all study participants.

REFEREES COMMENTS

Reviewer: 1

Recommendation:

Comments:
For authors

There is a need for studies that compare epidural steroids with conservative treatments. This study is timely and well designed. Its results are generalizable and of relevance for a broad specialty spectrum, including primary care as source of referral for interventional practice.

Abstract
Please explain the abbreviation ESI.
It would be useful to have a short explanation of "positive successful outcome".
In the conclusions, I would change the term "small" to "moderate". See also my comments for the results and discussion sections.
The last sentence is obscure for those who read only the abstract. I suggest deleting it.

We have made all of these changes, except for the comment about changing "small" to "moderate". In view of the Editors' comments about non-significant differences being considered negative, we have changed it to "modest" instead.

Methods
In the section Participants & Settings, the sentence "or > 3/10 if greater or equal to back pain" is unclear to me. Please re-write.

We have changed that to "≥ 3/10 if back pain is < 4/10". Because this was a comparative-effectiveness, we tried to be as inclusive as possible. We included patients with 3/10 pain if that was their main pain

(many service members tend to be stoic when reporting their pain) because pain is always subjective and in these patients, (despite sometimes low pain scores) it was bad enough to warrant them to seek treatment (and they would receive the same treatments as if there pain was 4/10).

At the end of the same section, please specify what psychiatric or medical conditions were considered as exclusion criteria; it is not intuitive which specific conditions "might preclude an optimal response to treatment", and why. At least some examples would help.

We provided examples since this was ultimately up to the provider. For medical conditions, we included poorly controlled diabetes (because of the steroids) or unstable angina; for psychiatric, we provided posttraumatic stress disorder or uncontrolled depression as examples.

On page 9, please specify how "moderate or severe canal stenosis" was defined.

< 12 mm (noted)

Results and Discussion
It seems to me that the results and discussion present too negatively the effects of epidurals. It is true that some of the outcomes, including the main endpoint, were either not significant of revealed modest effects. However, for some of them the effect was not trivial. For instance 51.6% vs. 33.8% of patients felt globally improved at three months. This seems to me a substantial difference in the frame of chronic pain treatment, where we hardly expect considerable improvements in outcome. Also the effect on worst pain was negligible. While I appreciate the conservative attitude of the authors, I would suggest reconsidering the wording throughout the manuscript.

Thank you. Whereas I believe that many of the authors agree with you, the Editors have specifically requested (see above) that we not interpret non-significant findings as "positive".

Additional Questions:
Please enter your name: Michele Curatolo

Job Title: Prof.

Institution: University of Washington

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may

in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF='http://www.bmj.com/sites/default/files/attachments/resources/2011/07/bmjpolicyondeclarationofinterestsmarch2014.pdf'target='_new'>(please see BMJ policy) please declare them here:

Reviewer: 2

Recommendation:

Comments:
Positive

The authors have performed a very elegant comparative-effectiveness study comparing ESIs to the first-line medication gabapentin for managing neuropathic low back pain. As a pain physician who manages patients with neuropathic low back pain, I think this gives us very informative and important messages in ways that will helpimprove clinical management. The article will be of interest to a wide range of medical doctors surgeons, internists and family doctors, and all of the different types of doctors who practice pain, so this matches the BMJ's mission to publish clinically relevant articles as well.

Radiculopathy is an enormous societal burden throughout all over the world, and the use of ESIs and medications such as gabapentin have both come under scrutiny because of their increasing utilization. This therefore seeks to address an extremely very important question. This study does not measure efficacy, but may be more relevant because there have been over 50 studies comparing anticonvulsants and ESI to placebo, and it's unlikely that the next study will provide a definitive answer. Rather, this is the question that primary care (and pain) doctors would like to be answered.

Based on recent literature, this study may be more clinically relevant because the authors elected to use a "true placebo" (intramuscular injection), which seems more difficult to blind, but unlikely to provide any benefit for patients with both spinal stenosis and herniated disc.

An interesting thing that the authors did was to measure blinding at 2 different stages (those who receive ESI are at risk for "unblinding" early, while those who receive a medication are at risk later because of side effects). Even though there was a trend towards better insight into treatment assignment for the gabapentin group at one month, this provides information on the assessment of "blinding" and how to maintain it successfully. Industry sponsored studies approved by regulatory bodies generally do not assess blinding, and the observation that more people who received gabapentin guessed their treatment group at 1-month compared to gabapentin at 1-day or ESI at 1-month raises the question about whether some of these other studies were actually blinded.

Thank you. We were actually more concerned about the effectiveness of blinding the epidural group (hence, assessing blinding just after the procedure). Most people now agree that only an IM injection is a true placebo (epidural local anesthetic is not), and our IRBs both wanted us to use a true placebo. Since

patients may sometimes feel sensations in their leg during TF epidural injections, we went to great extents (briefings etc.) to prevent this from happening. Very few gabapentin studies have assessed the efficacy of blinding (the FDA did not require it for the studies leading to approval), so we did not anticipate that some patients would correctly guess the group. We agree that this is important information.

Evaluating the primary outcome at one month by comparing a "Single ESI" to "oral gabapentin" might not be long enough to find any clinical differences between groups. As the authors suggested in their discussion section, few studies have examined the long-term effectiveness of gabapentin, but those that have indicate that the beneficial effects for neuropathic pain are most pronounced early on during treatment, which is similar with ESIs.

We agree that the beneficial effects are most pronounced early on in the course of treatment. In a meta-analysis presented before the FDA, ESI were shown to clearly provide benefit at 1-month, continue to provide some benefit at 6 weeks, but there is little difference at 3-months compared to a control injection. It was necessary to evaluate the primary outcome measure at 1-month because our IRB would not permit us to keep patients in a study without offering them "other treatments" who failed to benefit (i.e. we had to allow them to exit). In the limitations section of the discussion, 1st para, we added, "… the primary outcome being measured at 1-month (which was necessary because we allowed those with an unsuccessful outcome to seek other treatments)…"

The broad inclusion criteria and 3 types of hospitals in this study increase generalization and are consistent with guidelines on comparative-effectiveness research.

Negative

Several limitations that the authors addressed in their discussion section should be considered.

This study was powered to detect a difference between the 2 groups, but the numbers are probably not sufficient to detect subgroup analyses such as the type ESI done, age, or pathology.

We have noted this in the limitations section (added, "A larger study would be needed to determine whether certain patients (e.g. herniated disc vs. spinal stenosis) or treatments (transforaminal vs. interlaminar ESI) experienced better outcomes than others." We could have recruited a very homogenous population (as is recommended for a placebo-controlled study designed to determine efficacy), but this is not recommended for comparative-effectiveness studies where one wants it to be generalizable.

Another limitation that either needs to be noted is that the authors did not note the numbers of patients who underwent surgery. In addition, did soldiers or veterans have any different outcomes than civilians?

We have added data on surgery in table 2, which reveals no difference between the two groups for this outcome. The outcomes as stratified by military status was hidden in the section "Factors associated with…"; however, we added that there were also no differences for non-active duty veterans (there were less of those patients).

Minor

Methods (Participants & Settings)

Did all patients included in the study have leg ≥ low back pain? Were there any patients whose average leg pain score ≥ 4/10 on a NRS pain score but back pain was more severe?

There were 25 such patients, but of course all of these patients would have received ESI anyway. Because herniated discs nearly always occur in discs that are already degenerated (which can cause back pain; Iencean SM. Acta Neurochir (Wien) 2000), there are very few people who have leg pain without LBP.

Did all patients have MRI images before the study? The authors wrote in one of their articles that there is difficulty in establishing a cause-effect relationship between pathology and pain itself in that many patients have pathology in the absence of pain, so that ESI might show better effectiveness in patients who had signs and/or symptoms of neuropathic low back pain with concordant MRI imaging.

Yes, all patients had MRI before enrollment as concordant pathology was an inclusion criterion (see page 6- we elaborated on this).

Was there any reason to exclude subjects with ESI within the past 3 years? Did you exclude those subjects for recall bias or because you already knew how they responded?

This is a good question. We excluded them because we felt that it would be close enough in time for patients to remember what it feels like to have an ESI (it could affect blinding) and because they would be likely to respond similarly to a recent ESI.

Did patients undergo only single level ESI? Did you exclude patients if they had multi-level unilateral symptoms?

Patients only had a single-level ESI because the use of 2 injections during 1 visit is not validated and has never been shown to be more effective. In addition, contrast generally spreads to multiple levels (Weil et al. 2008; Botwin et al. 2004; Jeong et al. 2013; Goel et al. 2006; Stojanovic et al. 2002), and even at a single level transforaminal injection, the contrast will cover both the exiting and traversing nerve roots. In addition, over 40% of patients with lumbosacral radiculopathy have > 1 nerve root involvement (Czyrny and Lawrence 1996). We have added that patients with multi-level symptoms could be included on page 6 of the methods.

Randomization and Interventions

Regarding the allocation ("Allocation was performed in groups of 36 (Walter Reed and Johns Hopkins) or 18 by research nurses,"), would you explain the allocation method more in detail?

Johns Hopkins and Walter Reed were expected to enroll more patients because they see more patients, but also because of budgetary concerns (other sites were paid "per patient" so that they were limited as to how many patients they could enroll. Although we discussed the trial with patients beforehand, we did not allocate them to treatment groups until they came in for the procedure. This decreases the chance that the balance becomes skewed. We had a previous study in which subjects were allocated in advance of surgery, but some ended up being cancelled and this led to an unbalanced randomization process (we learned from our mistake). We try to allocate in blocks of high numbers so that investigators can't "guess" which treatment someone will be in. We wrote the reason we used high block numbers, why group allocation was larger at Walter Reed and Hopkins and added that allocation was stratified by study site (see page 7).

I realize that the latest guidelines that have been published online, presented and will probably also be presented at the U.S. Food and Drug Administration meeting this month on ESI state that transforaminal

ESI with particulate steroids should not be a first-line treatment. However, the authors should still note why they didn't perform bilateral transforaminal ESI for patients with bilateral pain in lieu of interlaminar ESI since it may be superior in cases with scar tissue.

There are no clinical trials examining bilateral transforaminal ESI for radiculopathy (there is one for axial back pain associated with spinal stenosis; Lee et al. 2009). At the FDA meeting, the data presented suggested that serious complications from transforaminal ESI result not just from particulate steroids, but possibly from the route itself (paralysis has been reported to the FDA with transforaminal dexamethasone). Because it is not common practice, these patients were for the most part "injection-naïve", there are no trials to support it, and it carries higher risk, we elected to perform a regular interlaminar ESI first. However, we think this could be a reasonable approach in patients with bilateral pain who fail interlaminar ESI.

Epidural Injection

Please explain why you used the volume of 4 ml for interlaminar ESI?

We decided on the volumes after examining all of the clinical trials that have been done (the volumes were in the middle of what was used).

Sham Injections & Maintenance of Blinding

Were subjects in ESI group also shielded from the image screen?

That subjects were visually shielded from the image screen is noted in the first sentence under "sham injections and maintenance of blinding" on page 8.

Regarding "sham injection," wasn't there any resistance when 3 ml of LA was injected into muscles or ligaments? Were there any patients who complained about discomfort during the injection? If the patient in sham injection group had had ESI (especially transforaminal approach) even once before, I would guess that a few of them could detect the difference between the true ESI and the sham procedure, though blinding was successful.

Please see our response to your earlier comment. Yes, some patients experienced discomfort from the intramuscular LA injections (though people do these frequently as "trigger point injections", just as those patients who underwent true ESI sometimes experienced discomfort. We actually had a "briefing" regarding maintenance of blinding to try to make sure it was properly implemented. Your comment regarding those who have had previous ESI being able to tell the difference is very true, which is why we did not allow anyone who had had an ESI within 3 years to enroll. Fortunately, blinding was successful right after the procedure.

Pharmacotherapy

I could guess that some patients who already had taken gabapentin before might detect that their medication was placebo.

The studies evaluating gabapentin generally did not evaluate the effectiveness of blinding so we assumed that blinding would be effective. Our exclusion criteria did not permit patients who had failed a previous

trial with gabapentin or pregabalin to enroll (see page 6).

Statistical Analysis

Regarding sample size and power analysis of the study. Please suggest a reference or explain why you used "a 1.0-point difference in pain scores between groups and SD of each group of 2.0."

The Editors asked the same question, so it is obviously an important one. Please see our response to their question. In the IMMPACT guidelines, they state that although a 30% or 2-point decrease in pain should be considered clinically meaningful, this does not mean that there should be the 2-point difference between a treatment and placebo in a clinical trial. In fact, in the only drug in the U.S. approved for chronic low back pain, the 2 studies that led to FDA approval comparing duloxetine to placebo found the difference to be $\leq$ 1-point (it looks to be about 0.75 in the first study for the high doses and around 1.05 or 1.1 points in the second study; Skljarevski et al. 2009, 2010). We chose the SD of 2.0 based on actual data from our previous studies.

Results

Outcomes

Primary care doctors and pain physicians could wonder how long it took to get to the final dose of gabapentin in the Sham ESI group, and whether the dose of gabapentin was related to the outcome in the binary logistic regression.

Please consider showing the results of factors associated with outcome by logistic regression as a table.

The dose titration schedule is noted under "Pharmacotherapy" in the "Methods" section, which indicated that patients were expected to be at the therapeutic dose of medication at least 5 days prior to their 1-month visit. An analysis of the gabapentin dose in the binary logistic regression revealed that there was lack of significance for an association with adverse outcomes (P = 0.06), as well as GPE or overall outcomes at 1 or 3 months. No association between gabapentin dose and other outcomes was found.

Given that only one factor was associated with an outcome in reanalysis of the data, we opted to include information related to their significance within the results section of the manuscript rather than another table.

End

Additional Questions:
Please enter your name: Jee Youn Moon

Job Title: Clinical Assistant Professor

Institution: Seoul National University Hospital

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF='http://www.bmj.com/sites/default/files/attachments/resources/2011/07/bmjpolicyondeclarationofinterestsmarch2014.pdf'target='_new'>(please see BMJ policy) please declare them here:

Reviewer: 3

Recommendation:

Comments:
Overall, this is a well-designed and well-written study. There are minor changes and clarifications that would be helpful in strengthening the study.

1. Please display the mean or median gabapentin dose.

The mean gabapentin dose per day was previously included in Table 3 under "Medication (dose) mg".

2. The authors should mention an additional heterogeneous characteristic of their inclusion criteria as a limitation: both radicular pain and "neurogenic claudication" were included, two different pathologic processes (defined by symptoms and distinct from radiologic diagnosis of disc herniation vs. stenosis), which many each respond differently to the interventions tested here.

In the limitations section of the discussion, we added the following: "A final limitation is our broad inclusion criteria which included patients on opioids, and those with herniated disc and spinal stenosis. These conditions are characterized by slightly different pathophysiological mechanism and may have different natural outcomes. A larger study would be needed to determine whether certain patients (e.g. herniated disc vs. spinal stenosis) or treatments (transforaminal vs. interlaminar ESI) experienced better outcomes than others."

3. What is the rationale for this outcome definition: "Reduction in analgesic medications corresponds to > 20% reduction in opioid use or complete cessation of non-opioid analgesic." Is 20% based on prior literature or arbitrary?

The First Author went to the FDA on January 13, 2015 and we discussed this issue at length for an upcoming clinical trial. There is unfortunately no set standard for defining clinically meaningful pain

reduction (or increase), though there is a strong consensus that there needs to be a "cutoff" or threshold. The reason for this is that most people take "as needed" pain medications so there is a lot of variability at baseline, and that people cannot just "stop" opioids, even if they experience improvement. For all other outcome variables (pain score, functional capacity, emotional functioning- see REFS), there are specific cutoffs that have been proposed as "clinically meaningful"(see Dworkin et al. J Pain 2008; Davidson M, Keating JL. 2002; Davidson M, Keating J. 2005; Devine J et al. 2011; Farrar JT et al. 2000; Jensen MP et al. 1994; Slater MA et al. 1997). We have used this cutoff for medication decrease for many studies (Cohen et al. Anesthesiology 2008, Anesthesiology BMJ 2008, Anesthesiology 2010, Anesthesiology 2014; Cohen et al. Ann Intern Med 2012; Cohen et al. Arch Intern Med 2011, many others), and there are other studies that have subsequently used the same ones we have.

4. "Positive" Composite outcome is based on a weak definition of a positive pain response. This definition of "a change greater than 2) that is below the MCID for leg pain according to some literature (3 points on the NRS scale). It would be worth adding additional categorical "responder" analysis using a more robust definition of clinically significant improvement in pain. The proportion of individuals who experience >50% pain reduction is commonly used in the literature as such a threshold. Using responder analysis may also possibly unmask larger differences between the two study groups.

Although there is some variability in what this cutoff is, we believe that the most commonly used cutoff is 30% (Farrar J et al. Pain 2001; Dworkin R et al. J Pain 2008 (IMMPACT guidelines). In both of these, they specifically mention a 2-point decrease on a 0-10 pain rating scale (which makes sense because it corresponds better to a 30% decrease in pain). Regardless, whether a 30% or 50% responder analysis was performed using average leg pain scores, which was the primary outcome for the study, no significant differences were found between injection and control groups.

5. The study would be more informative if outcomes were additionally stratified by primary pathology (stenosis vs. disc herniation) at the spinal level thought to be generating symptoms.
6. The study would also be more informative if results were stratified by bilateral ESI vs. unilateral ESI (TF vs. IL).

This information (etiology and injection type) was buried in the "Factors associated with outcome…". When we stratified ESI outcomes by level, performing an S1 TFESI was associated with a better outcome. However, we added some information to specifically highlight what was evaluated: "Otherwise, no associations were found among measures for pain, disability, medication reduction, global perceived effect, or composite outcome at either 1 or 3 months based on opioid use, military status, etiology (e.g. stenosis vs. herniated disc), pain duration >= 3 months, injection type (i.e. transforaminal ESI for unilateral pain vs. interlaminar ESI for bilateral pain), or dose of gabapentin."

Additional Questions:
Please enter your name: Zachary McCormick

Job Title: Physician

Institution: Northwestern University/The Rehabilitation Institute of Chicago

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?:

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF='http://www.bmj.com/sites/default/files/attachments/resources/2011/07/bmjpolicyondeclarationofinterestsmarch2014.pdf'target='_new'>(please see BMJ policy) please declare them here:

Reviewer: 4

Recommendation:

Comments:
I find this a slightly confusing read which is due, in part, to inconsistent reporting, there are also some inconsistencies in outcomes between the registry entry, protocol and the paper.

Title says low back pain, yet primary outcome is leg pain – which I don't understand.

There is some confusion as exactly what "neuropathic back pain " means (i.e. is it back pain accompanied by nerve root involvement or can axial back pain caused by nerve ingrowth into discs be included?). Therefore, we have changed it to the more generic and widely recognized term "sciatica".

The trial registry entry states

Primary outcomes as leg pain @ 1 and 3 months
Secondary outcomes as Back @ 1 and 3 months, ODI @ 1 and 3 months and Satisfaction @ 1 and 3 months.

Whilst, the paper states

Primary outcome as leg pain @ 1 month
Secondary outcomes as worse leg pain over the past week, average and worse back pain, ODI, adverse events, complications, reduction in analgesic medications, and global perceived effort. So a number of outcomes seem not to be pre-specified, and specific time points not indicated in the paper - though clearly in Table 2, interest is in 1 and 3 months (sample size is based on outcomes at 1 month). The

protocol has leg pain as primary and back (average and worse), OWI, medication usage, satisfaction, side-effects and complications.

Table 2 also introduces another unspecified outcome 'composite outcome'. The authors should clarify the inconsistencies, both within in the paper and also so that they tie together what was pre-specified in the trial registry entry.

Sample size calculation: I can see no rationale of a 1 point difference in pain. The assumptions on which the sample size is based are not referenced, more information on the assumptions that were used to derive the sample size are required, including how/why these values were chosen. Some of the assumptions (though not where these values were obtained from) are reported in the protocol, but should be included in the paper.

We apologize for the confusion but assure you that all outcomes were pre-defined. Registering trials is important but not always straightforward. The protocols (including power analysis) and manuscript are correct. The primary outcome measure is average leg pain @ 1-month. It could not be at 1 and 3 months because our protocol permitted people who failed to benefit to exit the study at 1-month. We have specified in the section "Co-Interventions, Outcome Measures, Follow-up & Missing Data" what the primary outcome measure was, and that it was pre-designated. This is noted in the protocol (section 6.3.2). We HAD to have a predesignated categorical outcome because (as is noted in the protocol), people with an unsuccessful outcome were allowed to exit the study to seek alternative care.

In the sample size section the authors also conduct post-hoc power calculation – I'm not quite sure why. This is never a good idea (large body of methodological and applied literature on this) and this calculation should be omitted from the paper.

Thank you for this comment. The sample size section for post-hoc power calculations has been removed.

The description of the randomisation is brief, just stating participants were randomised 1:1 using computer generated randomisation tables. More information required. The abstract states randomisation was stratified by site...this is not mentioned in the methods of the paper. The protocol describes this in more detail, and it should also be summarised in the paper.

Thank you for letting us provide additional information regarding randomization. To better address these questions, Johns Hopkins and Walter Reed were expected to enroll more patients because they see more patients, but also because of budgetary concerns (other sites were paid "per patient" so that they were limited as to how many patients they could enroll. Although we discussed the trial with patients beforehand, we did not allocate them to treatment groups until they came in for the procedure. This decreases the chance that the balance becomes skewed. We had a previous study in which subjects were allocated in advance of surgery, but some ended up being cancelled and this led to an unbalanced randomization process (we learned from our mistake). We try to allocate in blocks of high numbers so that investigators can't "guess" which treatment someone will be in. We wrote why group allocation was larger at Walter Reed and Hopkins on page 7 and added that allocation was stratified by study site.

Eligibility criteria - registry entry states participants were eligible if they were aged 17 or older. In the paper (and protocol) this is 18 or older.

The protocol and manuscript are correct. We had left open the possibility of enrolling patients who were 17 at certain military treatment sites since 17-year-olds can join the military. There were no 17-year-olds who were enrolled.

Abstract reports within arm differences from baseline, ok, but I would like the between arm comparison - which is reported in the paper. The results (in the abstract) on worse leg pain is an unspecified post-hoc outcome (which happens to be 'significant').

Thank you for this suggestion. The abstract reported "No significant differences were noted for the primary outcome measure (mean change from baseline -2.2 ± 2.4 in the ESI group vs. -1.7 ± 2.6 in the gabapentin group", and the "adjusted difference [95% CI], 0.5 [-0.2 to 1.3];" has now been included in this section as well. Additional clarification is provided in the results of the abstract and report secondary measures.

Treatment effect was analysed using ANCOVA. Did they adjust for anything apart from the baseline value of the outcome and sex (which was a post-hoc decision due to imbalance in the randomisation). Arguably, if sex was an important prognostic factor that balance was important then this should've been included in the randomisation procedure and/or pre-specified in the analysis plan. No ESI study has ever suballocated patients by gender, but this is generally controlled for in ANCOVA.

You are correct, the treatment effect was analyzed using ANCOVA and adjusted by baseline value and then sex. We did not specify in this protocol which variables were predetermined to be potential prognostic factors, but in all of our previous randomized controlled trials evaluating ESI (Ann Intern Med 2012; Arch Intern Med/JAMA Intern Med 2011; Anesthesiology 2014) gender was pre-specified to be an important factor and adjusted for (as well it should as there are significant differences between the sexes with regard to how pain is perceived). That said, even when we have considered gender to be an important factor, we don't generally sub-randomize based on gender as randomization should take care of this (I don't believe any RCT evaluating ESI has done this). But given the imbalance in gender in this study, a separate analysis was also conducted adjusting by gender.

The protocol mentions using Bonferroni corrections for multiple testing – yet there is (as far as I can see) no mention of this.

Thank you for bringing this to our attention. The statistical analysis section of the methods now includes language related to Bonferroni corrections for post-hoc testing, which was applied to all subgroup analyses. For example, examination of military status (non-service member, enlisted personnel, or officers) used a corrected P value of 0.05 / 3 = 0.016 as the cutoff for significance.

True ITT would use all patients randomised, there is some loss to follow up in the ESI group, where there follow-up outcomes could've been imputed. The protocol mentions imputation for replacing missing data may be used.

The manner in which missing data is handled is an important topic. Follow up outcomes for patients who exited the study were handled using the "last-observation-carried-forward" method, which may underestimate effect sizes when "dropouts" are due to lack of efficacy (Lane Pharm Stat 2008; Prakash et al. Int J Clin Pract 2008)." Data for missing values for participants was carried forward from their last observation. For patients who only presented with baseline data, pain and functional scores were carried forward at the 1 and 3 month time period. Other missing data at the 3-month time period was carried forward from values at the 1 month time period. Given these changes, the results including those listed in the abstract, outcomes, the post-hoc analysis, as well as Table 2 have been updated to reflect this information.

typo in Table 1. '3 to 12 months' - 26/73 is 35.6% (not 25.6%)

The table has been changed to reflect this correction.

Would like CONSORT flow diagram to resemble a recommended CONSORT diagram. E.g. How many received randomised treatment, how many were analysed etc.

We have changed the flow diagram to better illustrate the flow of subjects.

In the Discussion, they state as a limitation that it was a small trial – how can it be small if the sample size was achieved?

We have removed this in the first sentence in the limitations paragraph of the discussion.