

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form ([see an example](#)) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	Clinical decision making in spinal fusion for chronic low back pain: Results of a nationwide survey among spine surgeon
<b>AUTHORS</b>	Paul Willems, Rob de Bie, Cumhuri Öner, René Castelein and Marinus de Kleuver

### VERSION 1 - REVIEW

<b>REVIEWER</b>	<p>Peter Fritzell MD, PhD Spine surgeon, Orthopaedic department Falun Hospital 791 82 Falun Sweden</p> <p>I was recently invited to attend the dissertation of the main author of this paper. I do not consider this as a relevant objection with regard to my objectivity, but I leave this to the Editors</p>
<b>REVIEW RETURNED</b>	23/10/2011

<b>THE STUDY</b>	<p>1. This is a survey among physicians. No patients involved</p> <p>The STROBE check list was considered relevant for this survey among spine surgeons</p>
<b>RESULTS &amp; CONCLUSIONS</b>	<p>As this is a review among Dutch spine surgeons (orthopaedic and neurosurgeons), it should be stressed that the conclusions may not be representative for spine surgeons in other countries, and the authors indeed adequately address this issue in the Discussion section.</p> <p>However, there are many articles published in international journals during the last ten years focusing on the lack of consensus with regard to especially surgical treatment for CLBP, why I tend to consider the authors findings/conclusions as valid also in other countries</p> <p>Why is Tab 3a and 3b different with respect to included variables?</p>
<b>REPORTING AND ETHICS</b>	I regard STROBE as equivalent with CONSORT here
<b>GENERAL COMMENTS</b>	<p>The authors have addressed a very important issue, which I will comment as follows:</p> <p>During the last two decades, numerous studies and reviews are reporting on the importance of agreeing on relevant prognostic factors and diagnostic tests in order to select the right patient with chronic low back pain (CLBP) for the right surgical procedure. In two recent and comprehensive review articles in Spine this question is addressed: Cheng, J et al. "Clinical guidelines and payer policies on fusion for the treatment of chronic low back pain". 2011 Oct 1;36(21 Suppl):S144-63, and Wood et al. "Effectiveness of spinal fusion versus structured rehabilitation in chronic low back pain patients with and without isthmic spondylolisthesis: a systematic review". Spine.</p>

2011 Oct 1;36(21 Suppl):S110-9. The conclusion is the same as in the current survey; there is a lack of consensus.

This lack of consensus partly depends on the complexity of the CLBP condition as such, and in spite of all theories what causes the pain in the individual patient, the reason for this is in many cases not obvious even to the most experienced spine surgeon, or indeed any physician. The consequence may be that patients who should perhaps not have been operated on are getting surgery and vice versa, and also that a sub-optimal surgical procedure may be used in an individual patient.

To select the right patient for the right procedure is of course the optimal goal for each treatment advocated by any physician, but this becomes even more important when the treatment in question is "irreversible" which may in many cases, but perhaps not always, be the case with a surgical intervention.

In this light, the current study is important, and could serve as a platform for the continuing search aiming to identify relevant prognostic factors and diagnostics tests in order to select the right patient for the right procedure, be it surgical or not. It is of utmost importance that we as a profession agree on a "gold standard" both when it comes to select patients for surgery, and also what procedure should be used in the specific case. This will not be easy, and although each patient is an individual, it should be in everybody's interest to define a patient selection protocol as well as a surgical procedure protocol where most patients with a specific pathology/clinic may fit in, and only when this is not the case other solutions should be suggested.

The strategy of defining "gold standards" will make it important for the surgeon to specifically explain why a different solution is used which is outside the usual definition, a definition which, when it comes to surgery, could be described as the least invasive treatment with the least complication rate that has a comparably good chance of improving a specific low back pain condition, i.e. a treatment that can be defined as "cost-effective". In CLBP the word "improve" is crucial, as most therapies, including surgery, will not cure the patient, but making him or her better.

In order to his define "gold standard", an expression that should be used with great care, with regard to firstly select the right patient for surgery, secondly to select the right procedure, and finally to select the optimal rehabilitation program, the profession first need to agree that we currently do not agree. After this we should start working on finding the important predictive variables.

This is the more important as most procedures indeed will not cure the patient, meaning that there is a distinct risk that we as surgeons will conclude that a negative outcome after surgery may be due to the use of a less invasive surgical procedure (we have not done "everything possible"), and that it would have been better had we gone for a more technical demanding procedure, even to a price of a possibly higher complication rate. This is human, but a potentially dangerous way of thinking. Another potentially inadequate situation today is because we do not improve more than approximately 70% of our patients operated for CLBP, there is a continuous introduction of new implants on the market that are often used before the previous ones have been properly evaluated. This highlights the

	<p>importance of the message in the current manuscript.</p> <p>There are minor details that I miss in the manuscript. For example, it would be interesting to know the number of operations performed annually in Holland for CLBP. In Sweden the number of 100/100 000 inhabitants have been reported. The number could indeed say something about the indications for surgery, and thereby be of importance when it comes to comparisons between countries with regard to the view of prognostic factors and views on diagnostic tests. Also, is there a difference in indications between neurosurgeons and orthopaedic surgeons in Holland? Why do neurosurgeons have a seemingly more restrictive view on fusing patients? Finally, are there differences in outcome depending on who performs the operation – a neurosurgeon or an orthopaedic surgeon? However, I realize that this was not the primary scope of the article.</p> <p>The current study focuses on the urgent need for searching for predictive prognostic factors, including useful diagnostic tests, in order to agree on selecting the right patients with CLBP for surgical treatment. The important consequence will be a need for the , profession to discuss and agree on “gold standards”. The authors are to be complemented for their important work.</p>
--	---

<b>REVIEWER</b>	Jeremy Fairbank Nuffield Orthopaedic Centre, Orthopaedic surgery
<b>REVIEW RETURNED</b>	06/10/2011

<b>THE STUDY</b>	CONSORT checklist or similar is not appropriate. This is a study of surgeons, not patients
<b>RESULTS &amp; CONCLUSIONS</b>	I would like to see the evidence associated with each question better marshalled, to bring out deviations between surgeons and evidence
<b>GENERAL COMMENTS</b>	<p>Thank you for asking me to review this paper. The results of this survey are unsurprising and would probably be replicated in surveys of spine surgeons in all developed countries. It is important that this useful and well designed study is published. The reference list is to the most important material on predictive factors and rct's that have been done on spinal fusion for back pain patients. It does ignore the other main surgical intervention for back pain patients – disc replacement, where the same issues apply. Presumably some of these surgeons are doing disc replacements. Do they believe the same methods apply to these patients as well?</p> <p>I think the interpretation of these studies needs to be linked more clearly to the responses to each question. This is a matter of editing: “The literature says...”, “The surgeons say...” The literature and surgeons seem to coincide a bit but not all that well.</p> <p>You have distinguished between academic and non-academic surgeons. If you look at individuals who deviate from the literature, do they deviate consistently? In other words are there some surgeons who are well outside the ‘funnel plot’ of conventional practice?</p> <p>Turning to the interpretation of these findings, my view is that these findings reflect our profound ignorance about chronic back pain. I am of course influenced by the study that we ran (reference 11). We</p>

	<p>compared intensive rehabilitation with fusion and fusion was slightly better. These interventions are based on completely different conceptions of back pain and how to treat it.</p> <p>So where do we go from here? I think one problem is that all our studies are hopelessly underpowered. We need some large studies to explore the predictive capacity of these tests. I suspect we shall find that most are of little value. It is extraordinary that so few of us really assess depression and have ways of really understanding the impact of that assessment on outcome.</p> <p>Meantime knowledge is progressing rapidly on mechanisms of chronic pain and disc degeneration. This laboratory based knowledge has to be applied and tested in a clinical environment to try to predict outcome and select only the best candidates for this major surgery.</p>
--	--

### VERSION 1 – AUTHOR RESPONSE

First, I would like to thank the reviewers for their positive and valuable comments.

As suggested, the title has been shortened to (listed in red in the manuscript):

Clinical decision making in spinal fusion for chronic low back pain.

Results of a nationwide survey among spine surgeons.

We agree that for patients who undergo Total Disc Replacement (TDR) for chronic low back pain, probably the same issues in decision making apply as for patients who receive spinal fusion. The reason why we did not include issues on TDR in the present survey, is that in The Netherlands lumbar TDR is hardly performed at all. Four surgeons responded to perform TDR, only one of whom stated to perform more than 10 TDR/year. Therefore, we considered this group too small for analysis.

The surgeons' responses and findings from literature have been linked more clearly as you can see in the discussion on pages 10-12.

Indeed, we looked at academic and non-academic surgeons: the only significant differences we found, was that non-academic surgeons believed more in provocative discography (20 of 43 versus 3 of 18 academic surgeons,  $p=0.028$ , see p.8 of the manuscript) for decision making. Additionally, academic surgeons tended to operate less on patients below 30 ( $p=0.025$ ), this has been added in red in the results on p.8.

It is hard to verify whether there are some surgeons who deviated consistently from conventional practice or evidence from the literature. First of all, as there was very little consensus on all items, it is hard to speak of conventional practice and secondly, for many items there is no straightforward, generally acknowledged evidence.

Looking at subgroups (medical discipline, years of experience, academic or non-academic) who might have totally different opinions, there were no groups who deviated consistently on more than a few items.

(Reviewer 2)

We agree that the conclusions from the present survey among Dutch spine surgeons may not apply for all spine surgeons in different countries as was stated in the discussion.

In Table 3a the opinions of the surgeons on different prognostic tests have been grouped, whereas in Table 3b the actual use of these tests in practice was listed.

In this survey we focussed on 3 prognostic tests that are generally performed by the surgeons themselves. As facet joint blocks are mainly performed by pain physicians in the Netherlands we only included 1 item on this subject (question 20), whether the test was used for patient selection or not.

Unfortunately, the number of spinal fusions performed annually in The Netherlands for chronic low back pain, is not known. Implantation devices used for instrumented lumbar spinal fusions are registered but not stratified for different diagnoses. Presumably, like in Sweden we are talking about thousands of patients per year. This highlights the importance of central registration of spinal surgery, with which we hope to start next year in the Netherlands.

The differences we found between neurosurgeons and orthopaedic surgeons was that orthopaedic surgeons appeared to operate more on 3 levels or more ( $p=0.003$ ), tended to fuse more patients below 20 ( $p=0.063$ ) and believed more in discography ( $p=0.025$ ) as a tool in clinical decision making. These differences are reported in the results on p.8 of the manuscript.

Only for long tract fusions (3 or more levels) we found a more restrictive view on fusing patients for neurosurgeons. The reason for this cannot be extracted from the data of the present study. A possible explanation could be that orthopaedic surgeons are more familiar with long tract fusions from scoliosis surgery.

In the present survey we did not look at outcomes of fusion. It would be very interesting indeed, to look whether certain strategies of decision making provide better outcomes of surgery. Once again, spinal registration could be a tool to provide an answer to such questions.

#### **VERSION 2 – REVIEW**

<b>REVIEWER</b>	Jeremy Fairbank Nuffield Orthopaedic Centre, Orthopaedic surgery
<b>REVIEW RETURNED</b>	15/11/2011

<b>GENERAL COMMENTS</b>	I am happy with the changes
-------------------------	-----------------------------