The Appropriateness of Spinal Manipulation for Low Back Pain

By Paul Shekelle, MD, MPH

First I would like to tell the audience who I am and what RAND is and what role we play in this. I am an internist, and am on the faculty of UCLA. My research interests are in assessing the appropriateness and quality of health services, and I am the principal investigator of this project. The acronym RAND stands for research and development. RAND is a non-profit corporation performing research in the public good. RAND’s research programs include classified defense research for the military, applied economics, education, sociology, civil justice, and health sciences. The health sciences department is the largest non-university-based research center in the U.S. For the past decade, a central effort of the health sciences program has been in the development and application of methods to assess the appropriateness of medical procedures.

As all of you in this room can appreciate, trying to assess the appropriateness of the use of any medical procedure is a difficult task. Ideally, we’d like to assess appropriateness on the basis of scientific evidence of increased benefit relative to risk. However, there are relatively few clinical problems that have been definitively answered by a well-executed controlled clinical trial. Even for these clinical problems which have been examined this way, frequently the patient in your office differs in some crucial way from the patients enrolled in the research trial such that generalizing the results from that trial to your patient may not be valid. And, of course, there are many clinical problems that have yet to be examined in a scientific way.

On the other hand, there is the large body of knowledge on appropriateness contained in the collective minds of the physicians of this country. Frequently, however, recommendations for action based on this clinical knowledge are in conflict, even among nationally recognized experts in the field. Resolving these conflicts can be problematic.

Lastly, there is the problem of the infinite variability of patients, making it an impossible task to try and define in advance the appropriateness of any medical procedure for any given patient that may come through your office door.

What can we hope to accomplish from studies of appropriateness? I believe we can define specific clinical circumstances for the appropriate use of medical procedures for general, commonly recurring classes of
patients, that the clinician can use as a reference point, or guideline, to help him or her make the individual decision on appropriateness for the individual patient.

So how best to decide appropriateness? One way that has been developed, and I might add has been extensively used and is the method that has been the best studied, was developed at RAND and UCLA about 10 years ago, and has been undergoing improvement ever since. It involves a systematic literature review to capture all of the relevant literature; the development of clinically detailed indications describing patients who might undergo the procedure; the convening of a panel of experts; and the assessment of the indications by the experts for appropriateness, using a formal method designed to promote consensus, without forcing a lowest common denominator type of agreement. This method includes features of both the Delphi, namely that of anonymous ratings and multiple rounds of ratings with feedback of group results to individual members; and the Nominal group, namely the use of a roundtable, face-to-face discussion.

How did we use this method to assess the appropriateness of spinal manipulation for low back pain? First, we performed the literature review. We searched the Index Medicus and Medline for relevant articles, then searched the bibliographies of these articles for additional sources, and lastly had our sources reviewed by experts to ensure that we hadn’t left out any critical material. This search netted 74 sources, which included 21 controlled trials of spinal manipulation as a treatment for low back pain. From these sources, a 32-page review was prepared, which summarized what is known about utilization, complications, and efficacy.

We then developed a detailed list of indications for patients who could potentially undergo spinal manipulation. This meant describing patients with back pain in a systematic fashion, in as much detail as possible, using clinically important variables, without using so many variables that we would overload our panelists. We did this by observing clinicians in their offices, and questioning them about their clinical reasoning. We spoke with experts, and reviewed the literature. We attempted to explicitly define key terms. The result was over 1,500 separate descriptions of patients with back pain, categorizing each patient by the length of the symptoms, the presence of comorbid diseases, the clinical course of the pain, the history and response to treatment of previous episodes of back pain, the findings on physical exam, the findings on lumbosacral radiographs, and the findings on CT or MRI if these tests were performed.

We then select a panel of experts. These are the general criteria governing the selection of RAND’s expert panels. They have nine members; they are a mix of academic physicians and private practitioners; they include at least one member from each of the four major census regions of the country; and they include a
mix of those who do the procedure and those who do not. This last point we feel is especially important. No medical specialty sees all of the patients with a given clinical problem. For instance, all patients with coronary artery disease don’t end up in the cardiovascular surgeon’s office; many patients are seen by cardiologists, internists, and family practitioners. The input from these practitioners is important, as they may see a different spectrum of the disease than the cardiovascular surgeon does. This is also true for back pain, as patients seek care from orthopedic surgeons, internists, GPs, and chiropractors, among others.

Given these considerations, this is the composition of the expert panel selected: three chiropractors, two of whom are in private practice and one of whom is an academic physician; two orthopedists and one osteopath, all of whom are distinguished experts in low back pain; one internist, who is an expert in back pain and is the principal investigator on the Back Pain Outcome Assessment Team; one family practitioner, who is trained in manipulation; and one person with degrees in many disciplines, a past president of the North American Spine Society, who practices as a neurologist.

We mailed the literature review, the definitions, and the 1,500 indications to the panelists. In private, they assigned an appropriateness rating, on a scale from one to nine, with one being highly inappropriate and nine being highly appropriate, to each of the 1,500 indications. They then mailed the completed appropriateness rating forms back to us, and we created summaries. All of the panelists met at RAND in Santa Monica. They were provided with the summary ratings, as well as a reminder of their own individual rating. We then had an open discussion of each indication for appropriateness, concentrating on areas of disagreement. This discussion was quite lively at times. After each general class of indications was discussed, each panelist privately rerated each indication for appropriateness. These second round ratings constitute the final results of the process.

What were the results? We define AGREEMENT as occurring when, after discarding the one highest and lowest rating, the ratings of the remaining seven panelists cluster within any three point range, that is two, three or four. We define DISAGREEMENT as occurring when, after discarding the one highest and lowest rating, at least one of the remaining seven panelists assigned a rating in the highest tertile (meaning most appropriate), and one of the panelists assigned a rating in the lowest tertile (meaning most inappropriate). By these definitions, on the initial round of ratings there was agreement on 12 percent of the indications and disagreement on 37 percent, with 51 percent indeterminate, meaning the panel neither agreed nor disagreed. After the final round, there was agreement on 36 percent of the indications, with disagreement on 12 percent. The number of indications changed somewhat between the two rounds, as the panelists altered the
original indications to better fit clinical circumstances.

Comparing this panel to some previous RAND consensus panels. I find it interesting to note that this panel had the second lowest level of disagreement ever seen. For example, in a recently completed panel concerning the appropriateness of CABG and angioplasty, there was disagreement in the final round on 18 percent of the indications. The fact that a panel of orthopedists and chiropractors have less disagreement on the appropriate use of spinal manipulation for low back pain than a panel of cardiologists and cardiac surgeons have on the appropriate use of CABG and angioplasty is somewhat surprising.

From the nine point rating scale, we created three classes of appropriateness. We defined an indication as APPROPRIATE when after discarding the one highest and one lowest rating, the median of the remaining seven ratings was in the highest tertile, namely a seven, eight or nine, without disagreement. Analogously, we defined an indication as inappropriate when, after discarding the one highest and one lowest rating, the median of the remaining seven ratings was in the lowest tertile, that is a one, two or three without disagreement. Indications whose median rating was a four, five or six, and all indications with disagreement, were classified as equivocal, meaning they were neither clearly appropriate, nor clearly inappropriate.

Using these definitions, of the 1,550 final indications, 924 or 60 percent were classified as inappropriate, 112 or 7 percent were classified as appropriate, and 514 or 33 percent were classified as equivocal.

Which indications were appropriate, and which were inappropriate? Of course I can’t go through 1,500 different indications at this time, but I can tell you the general characteristics of each type of indication.

In general, the appropriate indications tended to be those patients with acute low back pain, defined here as pain of less than three weeks; those patients with none or minor neurologic findings, minor neurologic findings being defined as either a unilateral diminished ankle reflex; no evidence of sciatic nerve root irritation, defined as typical radicular pain and the presence of a positive ipsilateral straight leg raising sign; physical findings of joint dysfunction, which in general meant hypomobility; unremarkable lumbosacral radiographs; and no adverse response to prior spinal manipulation, if any.

The general characteristics of those indications felt to be inappropriate were the following: no response or an unfavorable response to prior spinal manipulation; the presence of contraindications on lumbar x-rays, these being such things as malignant tumors, osteomyelitis, inflammatory arthritis, septic arthritis, acute or
unhealed fracture, achondroplasia, and others; or no x-rays in the presence of risk factors for contraindications (such risk factors being fever, history of malignancy, severe osteoporosis, age greater than 50, significant trauma, and the like); pain of greater than six months duration, prior laminectomy (and this referred to manipulation within the area of prior laminectomy, not remote to it), major neurologic findings, these being the cauda equina syndrome, or progressive motor weakness; a lack of response to the current spinal manipulation; and the presence of a centrally herniated disc, spinal stenosis, or a large posterolateral herniated disc in the presence of sciatic nerve root irritation or any neurologic findings.

In the context of this discussion on the continued appropriateness of spinal manipulation in the face of a lack of response to current spinal manipulation, there was unanimous agreement on the following: An appropriate trial of spinal manipulation for patients with low back pain is two weeks of treatment for each of two different types of manipulation, after which, in the absence of improvement, continued manipulation is not indicated.

What were the characteristics of equivocal indications? Sciatic nerve irritation is the clinical factor responsible for the greatest number of equivocal indications with disagreement. The presence of a herniated disc on CT/MRI, especially in the absence of signs or symptoms other than back pain, was also a characteristic of many equivocal indications. The longer the patient had back pain, the more likely the indication was to be equivocal.

It is important to keep in mind that the percent of indications given a rating of "appropriate" does not translate into the percent of patients with back pain who might be appropriate for spinal manipulation. Patients with the cauda equina syndrome, or malignancies, or an abdominal aortic aneurysm, or who take coumadin are all felt to be inappropriate candidates for manipulation, while patients with uncomplicated acute low back pain are felt to be appropriate candidates. In terms of numbers, the inappropriate indications I just listed outnumber the appropriate indications by four to one, but my guess is that the average chiropractor sees many more patients with uncomplicated acute low back pain than patients with the cauda equina syndrome. We don’t accurately know what kinds of patients are presenting to clinicians’ offices and subsequently undergoing spinal manipulations. Future research is needed to answer this question.

Several important limitations deserve mention. The first is that we dealt only with spinal manipulation, and not the spectrum of conservative medical or chiropractic care. Such care may include other forms of treatment that may prove beneficial for patients with back pain. Secondly, the alternative treatment against
which spinal manipulation was being compared was not explicitly defined. Rather, each panelist was comparing spinal manipulation against his own internalized "best available alternative," and these alternatives may have differed between panelists. One panelist may have been comparing spinal manipulation with bedrest and analgesics, while another, for the same indication, may have been comparing it with surgery. In the future, we will need to be more explicit about alternative treatments. Lastly, patient preferences are not accounted for in this process. Such preferences can greatly affect the appropriateness of procedures. Some patients may prefer a trial of spinal manipulation rather than bedrest, while others may prefer the reverse.

To summarize, let me state that this conference was not the "final word" on the appropriate use of spinal manipulation for low back pain. Rather, it is a first step, in that for the first time, orthopedists, chiropractors, and others assessed the appropriateness of spinal manipulation using a well-established method developed for such assessment. Our results show that there are significant areas of agreement, for both appropriate and inappropriate indications. Additionally, areas of disagreement are clearly defined. Finally, a number of indications for which the panel felt that there was not enough scientific data or clinical experience to assess appropriateness have been identified, and these areas may be fruitful avenues for future research.

Paul Shekelle, M.D., M.P.H.
The RAND Corporation
Santa Monica, California

Editor’s note:

This research is a joint undertaking of RAND, and the UCLA Division of General Internal Medicine, Department of Medicine.

This series of studies is funded by the Consortium for Chiropractic Research (CCR) and the Foundation for Chiropractic Education and Research (FCER).

If you would like a copy of the first two stages of the RAND study, you may call or write the RAND Corporation Distribution Services, 1700 Main St., Santa Monica, CA 90407; (213) 393-0411. They will accept a check by mail or a credit card number over the phone.
"The Appropriateness of Spinal Manipulation for Low Back Pain: Project Overview and Literature Review" costs $4.00 plus .33 tax. "The Appropriateness of Spinal Manipulation for Low Back Pain: Indications and Ratings by a Multidisciplinary Expert Panel" is $7.50 plus .61 tax. First class postage is .75 per publication and discounts are available for orders of 25 or more.

The Journal of Chiropractic Technique will be publishing the opening and closing presentations and summaries of each module of the Sixth Annual Conference on Research and Education in the February, 1992 issue (vol. 4 no. 1).

For more information, please contact:

Williams and Wilkins
428 e. Preston Street
Baltimore, MD 21202
(301) 528-4000

Page printed from: