

NEWS

The Future of Vertebroplasty: Are Randomized Controlled Trials, or Clinical Experience, the More Appropriate Guide?

Neil A. Andrews

Managing Editor, IBMS BoneKEy

Between the time it was first introduced in France in 1984 and its extensive adoption worldwide through the first decade of the 21st century, percutaneous vertebroplasty (VP), a procedure where physicians insert cement through a needle into a patient's fractured vertebra to relieve pain and disability not effectively managed by more conservative treatment such as bed rest and use of narcotics, has lacked evidence from double-blind, randomized, controlled trials (RCTs) – until now. Two RCTs published together in the August 6, 2009 edition of the *New England Journal of Medicine* (1;2) both found that VP relieved pain and disability in patients with painful osteoporotic vertebral compression fractures. However, a sham procedure, used as the control and akin to the real thing except for the injection of cement, relieved pain to a similar extent such that there were no statistically significant differences in primary or secondary outcomes between the two groups.

Calls for RCTs have been made for years in the published literature on VP and have persisted recently, so certainly these two trials, both published in a prestigious medical journal, would be welcomed by the field for providing convincing data about a very popular medical procedure. To the contrary, instead of providing definitive evidence about VP's effectiveness or lack thereof, the new findings have stimulated a heated controversy. Some experts maintain that VPs should no longer be performed, while others say it is appropriate that the procedures be continued but only for research purposes, namely for patients enrolling in clinical trials. Meanwhile, other experts, primarily those with extensive clinical experience performing the procedure in thousands of patients, say that VP's benefits are so obvious and dramatic that to

deny patients VP, or to restrict its use, is entirely unacceptable. The essence of the debate between such advocates of VP and those who have grown much more skeptical because of the recent *NEJM* trials is this: do these new results from two RCTs now trump more than a decade's worth of physicians' clinical experience showing that VP produces striking reductions in pain and disability, alleviating intense suffering that had remained intractable despite all other accepted treatments?

History

The story of percutaneous VP begins in France, where the first one was performed by Pierre Galibert, an oral surgeon, in 1984. "At the beginning, percutaneous vertebroplasty was performed only in patients with vertebral hemangiomas," according to Hervé Deramond, a neuroradiologist who assisted Galibert during VP's debut. The technique was then extended to patients with pain resulting from tumors and osteoporotic fractures, though by then, in the late 1980s, it had not spread worldwide. "The technique at that time was confined mainly to Europe, and most of the teams performing the procedure were neuroradiologists in France," according to Dr. Deramond, who continues to publish articles on VP and balloon-assisted VP (kyphoplasty (KP)) while based at Amiens University Hospital in Amiens, France.

VP then crossed the Atlantic to arrive in the US in the early 1990s after Jacques Dion and Mary Jensen, then researchers at the University of Virginia, heard descriptions of the procedure at a meeting of the American Society of Neuroradiology, and began treating patients with tumors and osteoporosis. "People ask why we didn't go ahead and randomize patients at that time,

but we were doing the procedure on patients who had tumors, and patients with osteoporosis who were all transplant patients on steroids who could not get off of them and whose bones were fracturing,” explains Dr. Jensen, director of interventional neuroradiology and a professor of radiology and neurosurgery at UVA. “So I felt it was not ethical to randomize these groups of people for whom the procedure was essentially a last ditch effort to alleviate their pain, and many of them responded.”

Consequently, with the success of these early cases, Dr. Jensen and colleagues performed VP on patients with osteoporotic compression fractures whose pain was not sufficiently controlled by medical therapy. This culminated in the 1997 publication of a paper in the *American Journal of Neuroradiology* describing the technique and reporting results from the first 29 patients. “Once this paper came out, people started asking us how we did the procedure, and we passed it on from there,” Dr. Jensen says. VP has certainly caught on: an editorial (3) accompanying the two *NEJM* studies cites Medicare data showing that the number of VPs performed in the US from 2001 to 2006 more than doubled, from 4.3 to 8.9 procedures per 1000 people.

The Current Studies and Author Reactions

Despite the number of procedures – actually, *because* of the number of procedures – David Kallmes, lead author of one of the *NEJM* articles and, in fact, Dr. Jensen's fellow at the time the procedure was being developed at UVA, believed that a trial was warranted. “I thought that if we are going to be doing so many vertebroplasties a year, over time that adds up to a lot of patients, so we should really know [that the procedure works]” says Dr. Kallmes, a professor of radiology at Mayo Clinic in Rochester, Minnesota. Dr. Kallmes also had another reason to question VP. “I had long experience with the procedure and I had high confidence in its efficacy, but there was some uncertainty in my mind because it seemed that no matter how we did it, it always seemed to work, irrespective

of the volume [of cement used], or the operator, or any other parameter, which made me wonder.”

As lead author of the other *NEJM* study, Rachele Buchbinder was also looking for confirmation that VP works. “We decided to do the trial because there had been many before-and-after studies that suggested there was a very beneficial effect, and there was also some worry that vertebroplasty might actually increase the risk of subsequent vertebral fractures, so we decided that a randomized, controlled trial was definitely needed first to confirm the benefits that were seen, and also to look at the risk of adverse effects,” says Dr. Buchbinder, a rheumatologist and clinical epidemiologist at Monash University in Malvern, Australia.

In Dr. Buchbinder's study, 78 patients with one or 2 painful osteoporotic vertebral fractures were randomized to receive VP or a sham procedure. Patients who received VP did show reductions in the trial's primary outcome of overall pain, but so did patients in the sham group, such that there were no statistically significant differences between groups at 1 week, 1 month, 3 months, and 6 months, nor were there differences in secondary outcomes like quality of life, physical functioning and other pain scores. Dr. Buchbinder says she was not surprised by these findings. “There have been numerous examples of treatments that looked highly effective in inferior studies, and then you properly evaluate them [and find] they are no better,” she said. Dr. Buchbinder noted first-hand experience that she has to support such a claim, namely, a previous study of her own, corroborated by numerous additional negative trials, that found no benefit of shockwave therapy for plantar fasciitis. “Shockwave therapy receives lots of good press, lots of people believe in it, and we did a negative trial, and there have now been nine negative trials,” she said. Based on her results published in the *NEJM*, Dr. Buchbinder thinks VP is not an effective procedure. “I can't understand why anyone would want to have the procedure or why anyone would continue to do it,” she says.

In Dr. Kallmes' study, results similar to those from Dr. Buchbinder's trial are reported. Specifically, 131 patients with one to three painful osteoporotic vertebral compression fractures were randomized to receive either VP or a sham procedure. The trial found no statistically significant differences between the two groups, at one month, in any of the primary or secondary outcomes that were measured, including scores on a disability questionnaire and ratings of back pain by patients – both treatments improved those outcomes to a similar degree. “I was really, really surprised by the results,” Dr. Kallmes said. “I thought there would be some benefit [to injecting cement], but we didn't find any.” Notably, Dr. Kallmes does not say that VP doesn't work. In fact, he says that his trial has confirmed that it does work. “What we show, though, is that it works with or without the cement,” he says. Dr. Kallmes believes that it is acceptable for physicians to keep performing VPs, but only as part of clinical trials.

Jeffrey Jarvik, one of Dr. Kallmes' co-authors and a professor of radiology and neurosurgery at the University of Washington in Seattle, said that he too felt surprised, but not primarily by the trial's findings themselves. “I've been more surprised by the reaction. I firmly believed that there was equipoise, that a trial needed to be done, that the evidence was not out there that this was a really efficacious procedure, and so I had an open mind actually that it could go either way. However, the reaction to the trial has shown me that not everybody has an open mind,” he says. Dr. Jarvik says he has heard from patients, who say the procedure has saved their lives, as well as from physicians who have reported successful results in case series, and from representatives from industry who are also very favorable towards the procedure. “That anecdote and case series are powerful but potentially misleading, and that's why we do randomized trials, to avoid the issues of bias, and regression to the mean, and the placebo effect, and all of these things that can masquerade as a true effect when in fact there isn't one.” Like Dr. Kallmes, Dr. Jarvik believes it is appropriate for physicians to offer VPs to patients, but

only if the procedures are done in the context of clinical research studies.

The Critics Respond

Top physicians in the VP field who perform the procedures in very large numbers are unanimous in asserting that their clinical experience of success with VP outweighs the results from the two *NEJM* trials. “I don't think we can discount the over 15 years of data that we have gathered on vertebral augmentation, with hundreds of thousands of patients benefiting, based upon these two recent studies with small numbers of patients,” according to J. Kevin McGraw, a VP expert and section head of Riverside Interventional Consultants in Columbus, Ohio who learned VP from the UVA experts during his training there. Thinking along similar lines is Allan Brook, director of interventional neuroradiology at Montefiore Medical Center and an associate professor of clinical radiology and neurosurgery at Albert Einstein College of Medicine in the Bronx. “This is one of the best procedures, for pain relief, in any medical field. It really works, and it's very safe,” asserts Dr. Brook, who has performed thousands of VPs. “Patient selection is a key component. I've had patients who have been bedridden in the hospital for 3 weeks, in severe pain and on narcotics, and after you perform the procedure, they are walking the next day. That's not a placebo,” Dr. Brook says.

Dr. McGraw and Dr. Brook are critical of the notion that a medical procedure must have the seal of positive RCTs to be considered effective. “If you look at the practice of medicine, from a historical perspective, you don't see sham appendectomies, or sham gall bladder removal operations, and make comparisons to those who actually had their appendices or gall bladders removed,” Dr. McGraw says. “Medicine in and of itself does not lend itself to doing randomized, double-blind, controlled trials. Very few procedures are subjected to those rigors of the scientific community.” Dr. Brook expresses a similar sentiment: “Randomized controlled trials do not always simulate what comes into any doctor's office. The majority of the patients screened do not participate in studies. Good science is of course

important, along with experience in each subspecialty. Common sense and good clinical judgment make a good doctor.”

Advocates of VP do not just emphasize their clinical experience to justify their continued support of the procedure in the absence of positive RCTs, but they also point to what they view as damning flaws in the 2 *NEJM* RCTs themselves, weaknesses that prevent them from accepting the conclusion that VP does not work any better than a sham procedure. First, they say the two studies are too small to make any definitive judgments. As an example of this criticism, they point to a trend in Dr. Kallmes' and Dr. Jarvik's study showing that patients in the VP group exhibited a higher rate of clinically meaningful improvement in pain (64% of patients in the VP group versus 48% of patients in the sham group) at one month. The trend was not significant ($P=0.06$), but they say had the studies included more patients, such a trend would have become significant. Dr. Kallmes says that the trend is “encouraging” to him, but he does not place too much stock in it, nor does Dr. Jarvik, because clinically meaningful improvement in pain was not a pre-specified outcome of the study. Speaking more generally about the criticism that their study enrolled too few patients to reach decisive statements about VP's effectiveness, Dr. Kallmes, clearly exasperated by the charge, disagrees strongly. “We had great power to show small differences – 80 to 90% power to show the minimal clinically relevant difference [in pain and disability].”

A second criticism of both *NEJM* trials concerns patient selection: critics argue that the patients who declined to enroll in the studies were likely in much greater pain and at higher disability (and therefore did not want to potentially receive a sham procedure) than those who enrolled in the studies, and that had patients in worse shape at baseline actually been enrolled, the trials would have found statistically significant differences between groups. Dr. Kallmes does not think this is a valid criticism, and points to a study he did, published earlier this year in the *American Journal of Neuroradiology*, comparing baseline characteristics of patients who

enrolled in his trial versus patients who did not enroll. He found that the baseline back pain-specific disability of patients who didn't enroll was the same as that of patients who did enroll. In that study, Dr. Kallmes wrote that he and his co-authors were unable to directly compare baseline pain between those who enrolled and those who didn't because of differences in the way questions regarding pain were asked in the two groups. However, Fergus McKiernan, director of the Center for Bone Diseases at Marshfield Clinic in Wisconsin who has written extensively about VP and KP, dismisses the notion that patients in the two *NEJM* studies did not have enough pain at baseline. “If you look at the data, that's not a legitimate criticism. The average pain level in both of those trials is comparable to that observed in other studies.”

A third criticism concerns crossover data from Dr. Kallmes' study. Unlike Dr. Buchbinder's trial, Dr. Kallmes and his colleagues allowed patients to cross over from one arm of the study to the other if their pain persisted. Results showed that at 3 months, more patients in the control group (43%) crossed over to the VP group compared to the number of patients in the VP group who crossed over to the control group (12%), a difference that did reach statistical significance ($P<0.001$). The trial's critics view that result as evidence that VP was more effective than the sham procedure, but the study authors, who are still analyzing the crossover results, say there are several possible interpretations, and it is impossible at this point to know which one is accurate. For instance, while they note it is possible that there was in fact a difference between groups in a health-related outcome that they failed to measure, it is also possible that patients had become unblinded to what intervention they had received, and the unblinding, rather than a true difference in effectiveness, could explain the differential crossover.

While critics of the *NEJM* trials do acknowledge that the placebo effect, the phenomenon of regression to the mean, and the favorable natural history of vertebral fractures explain some of the benefits of VP, the magnitude and speed of improvement in

pain and disability they see in their patients suggests those potential confounding factors play only a relatively minor role. In addition, though the critics concede that nobody really knows for sure the mechanism by which VP assuages pain, they believe, as do most experts, that injecting cement into the fracture prevents it from moving, just as an external cast does for a broken arm or leg, and it is this stabilization of the fracture that produces pain relief. Finally, though they acknowledge that the jury is still out on whether VP increases the risk of fractures in vertebrae adjacent to those that receive VP, they insist that the procedure is safe when performed by an experienced physician.

Which Studies Should Come Next?

Critics who perform VPs as a routine part of their clinical practice bristle at a possible future where the procedures are performed only for research purposes. They suspect that if this is ultimately what happens, wealthy individuals will pay for VPs, those who can't afford them will be forced to enroll in clinical trials, and individuals living in rural areas far away from academic medical centers who cannot pay for VPs will be completely out of luck. Dr. Kallmes and Dr. Jarvik, however, are not the only ones comfortable with the assertion that VP should be conducted in the context of research in the belief that the procedure still merits further study. "It's a bit of hubris to think that 113 patients who underwent vertebroplasty in two studies can change the way we think. There are very few single studies in medicine that shape it forever," Dr. McKiernan says.

Dr. McKiernan has written extensively that more rigorous attention to patient selection is what future studies of VP require, a conclusion he has also reached after examining the 2 *NEJM* trials. As noted above, he says the problem is not that patients enrolled in the trials weren't experiencing enough pain pre-VP, but rather that studies of VP, including the current ones, have not been careful to distinguish vertebral fracture pain from other types of pain, such as postural fatiguing pain that patients often feel after suffering a vertebral fracture, or other back pain that co-

morbidities might cause. "It's important to make sure that patients and physicians understand that vertebroplasty only helps fracture pain. If you enroll people with ambiguous back pain generators and put cement in the vertebrae, it's not going to help," Dr. McKiernan explains. Similarly, Dr. McKiernan also argues that studies of VP have not suitably appreciated the reality that not all fractures are the same, in particular that some fractures move when a patient changes postures, by going from lying down to standing up, for instance. These fractures, which Dr. McKiernan and his colleagues have termed "dynamically mobile fractures," may have different outcomes after VP, something future studies will need to address, he says. Finally, Dr. McKiernan believes that future trials of VP should not include a sham-operated group since sham procedures are not offered to patients in actual medical practice.

Dr. Kallmes and Dr. Jarvik, for their part, are currently designing other studies they believe will bring clarity to the VP field, including a 3-arm trial that would include a VP group, a KP group, and a sham-operated group; a trial comparing their sham procedure, which could have an effect perhaps because it employs a local anesthetic, to a saline injection; and a trial investigating whether subgroups of patients with specific features apparent on baseline MR imaging of their fractures would benefit from VP.

Because of the results from the *NEJM* VP studies, studies of KP versus VP are indeed necessary, according to Richard Deyo, a professor of evidence-based family medicine at Oregon Health and Science University in Portland and a back pain expert. "We need comparative trials of vertebroplasty versus kyphoplasty, which is the alternative I think many people will turn to," he says. "Until we have equally rigorous trials there I think people will assume kyphoplasty is probably more effective than vertebroplasty." Currently, there are no published comparative trials of KP versus VP, and though there is one RCT showing a statistically significant advantage of KP over traditional medical management, critics of

that particular trial note that the study was not blinded.

A Shifting Playing Field

Will the new findings on VP result in more studies from other investigators in the field? Dr. Kallmes is optimistic. "Nobody was interested in doing research until now. Compared to the numbers of patients treated each year in the US, there is a disappointing number of real research trials going on, but I'm very hopeful that our results will stimulate more studies." Other experts hope that the new findings will also encourage regulators to reassess how they evaluate medical procedures like VP and KP. "If vertebroplasty or kyphoplasty were drugs, they would still be in phase II trials," says Nelson Watts, director of the University of Cincinnati Bone Health and Osteoporosis Center in Ohio. "Maybe what the FDA needs to do is to be a bit more rigorous with the measures they follow for procedures [like vertebroplasty and kyphoplasty]. Should this procedure have been let loose on the population without a randomized trial? That's something I think the FDA probably needs to re-consider," according to Dr. Watts, who nonetheless says he would still recommend VP to a subset of patients that would benefit from a quick return to mobility.

There is one other change that Dr. Kallmes and Dr. Jarvik think is warranted: now that 2 negative RCTs of VP have been published, they believe the playing field has shifted

such that staunch believers in the procedure must now step up: the onus is on them to support their assertion, with proper trials, that VP is effective. While Biblical miracles may be hard to support with evidence, Dr. Kallmes thinks a medical miracle – a procedure that firm advocates do tend to speak of in glowing terms – shouldn't be so hard to substantiate. Consequently, he offers a challenge to those who say that VP works:

"Go prove it," he says.

References:

1. Buchbinder R, Osborne RH, Ebeling PR, Wark JD, Mitchell P, Wriedt C, Graves S, Staples MP, Murphy B. A randomized trial of vertebroplasty for painful osteoporotic vertebral fractures. *N Engl J Med.* 2009 Aug 6;361(6):557-68.
2. Kallmes DF, Comstock BA, Heagerty PJ, Turner JA, Wilson DJ, Diamond TH, Edwards R, Gray LA, Stout L, Owen S, Hollingworth W, Ghdoke B, Annesley-Williams DJ, Ralston SH, Jarvik JG. A randomized trial of vertebroplasty for osteoporotic spinal fractures. *N Engl J Med.* 2009 Aug 6;361(6):569-79.
3. Weinstein JN. Balancing science and informed choice in decisions about vertebroplasty. *N Engl J Med.* 2009 Aug 6;361(6):619-21.