December 15, 2014

NASS Comments on Draft AHRQ Technology Assessment on Pain Management Injection Therapies for Low Back Pain

(http://www.ahrq.gov/research/findings/ta/index.html)

The North American Spine Society (NASS) appreciates the opportunity to comment on the AHRQ draft technology assessment, *Pain Management Injection Therapies for Low Back Pain*. NASS is a multispecialty medical organization dedicated to fostering the highest quality, evidence-based, ethical spine care by promoting education, research and advocacy. NASS is comprised of more than 8,000 physician and non-physician members from several disciplines, including orthopedic surgery, neurosurgery, physiatry, pain management, neurology, radiology, anesthesiology, research, physical therapy and other spine care professionals. Questions may be submitted to Pam Hayden, Director of Research & Quality Improvement at phayden@spine.org or 630.230.3690.

**METHODS**

NASS commends the authors for reviewing and synthesizing a large volume of literature. However, as specialists in spine care, we feel compelled to highlight a few points that contribute additional information on this topic with a primary focus on the poor methodology used for the review.

**Study Selection**

We understand that in strict adherence to the evidence-based process, randomized controlled trials (RCTs) were considered the highest level of evidence in this project. However, as an organization also deeply dedicated to EBM, we feel it is incumbent upon us to point out that although RCT’s are considered the gold standard in research, there are varying levels at which they are conducted. Including an assessment of the quality of RCTs as well as considering other well-done studies, such as well-designed observational studies, is imperative to the evidence-based process. For example, a level I study comparing nitrates to antacids for the treatment of nonspecific “chest pain” would have little value despite good study design. "Evidence based medicine is not restricted to randomized trials and meta-analyses. It involves tracking down the best external evidence with which to answer our clinical questions."1 We understand that these studies are very much a reflection of the quality of the literature base, however, well-designed and implemented studies that provide categorical data, as opposed to means of continuous data, on outcome measures including pain relief, functional outcomes, decreased use of other health care, surgery-sparing effects, and decreased use of opioids are needed to inform health care providers on which patient subgroups a given intervention may be effective.

NASS would encourage AHRQ to also consider observational studies. Concato found that "well-designed observational studies (with either a cohort or a case–control design) do not systematically overestimate the magnitude of the effects of treatment as compared with those in randomized, controlled trials on the same topic."2 Concato stated, "The popular belief that only randomized, controlled trials produce trustworthy results and that all observational studies are misleading does a disservice to patient care,
clinical investigation, and the education of healthcare professionals.” An evidence base comprised of well designed and implemented observational studies on consecutive patients can yield moderate to high quality evidence in accordance with GRADE. Unless multiple high quality RCTs with appropriately selected patients and technically accurate injections are available, observational studies should not be excluded from a comprehensive systematic review. There are multiple, methodologically rigorous studies which included large cohorts of consecutive subjects that offer additional insights into the clinical effectiveness of these injection procedures.

Specific to assessment of spinal injection therapies, many high-quality studies are excluded from this review that provide important evidence regarding the use of spinal interventions. For example, prospective observational studies show good short-term and long-term (one-year) outcomes for lumbar transforaminal epidural injections and lumbar facet joint injections. These findings are supported by very large retrospective studies with high-quality data, such as a Mayo Clinic study involving more than 2,000 subjects. These are just a few examples, and many other examples exist. These studies are inappropriately omitted from the AHRQ technology assessment. While the studies just cited involve patients with different low back symptoms, they all share one important feature. The study populations are well defined, not characterized by a symptom, such as back pain or sciatica; they have a radiographically confirmed patho-anatomic diagnosis that is responsible for their symptoms. This is a critical issue in the assessment of any study involving a targeted intervention.

In addition, it is imperative to recognize that study methodology is pointless if the procedures being assessed are not performed on appropriately selected patients using accurate technique. An RCT with sound randomization, excellent blinding, and no losses to follow-up is of no value if the patients did not have the condition and the procedure was not conducted accurately. Stratification of studies by acceptable, technical performance of the procedures is critically important. Other significant considerations regarding methodology include underlying diagnosis and its natural history, heterogeneity of procedural techniques along with use of imaging, and statistical analysis.

**Diagnostic Specificity**

Low back and radicular pain are merely symptoms, not diagnoses. Investigations of targeted injection therapies based on patients with a specific anatomic diagnosis repeatedly demonstrate high success rates for clinically meaningful changes in back pain and disability. Alternatively, spinal injections that treat back pain without a confirmed anatomic diagnosis yield poor results. This is an important distinction, however we feel that in this assessment diagnostic etiologies were inappropriately combined throughout the paper. Of the 29 studies comparing epidural steroid injections to placebo, 22 specified radicular pain alone, 6 included a mixture of radicular and back pain, and one study included patients with back pain alone. For perspective, imagine a hypothetical systematic review of prescription medication for the treatment of cough, a symptom. A few studies may show beneficial effects from antibiotics in a group of patients with bacterial pneumonia, a specific diagnosis, whereas pooled data from heterogeneous groups – including viral bronchitis, chemical pneumonitis, asthma, lung cancer, etc. – would produce different effects. If these pooled effects showed that many different medications had minimal impact on cough from various sources, would we abandon prescription antibiotics for pneumonia?

Additionally, the identification of the underlying etiologies of pain is essential as different pathologies not only have varying responses to treatment, but also have different natural histories. Thus the time
frame of follow-up to determine clinical utility becomes imperative. Some conditions, such as intervertebral disc herniation, can result in debilitating pain, but have an overall favorable natural history. This would be in contrast to spinal stenosis, which is less likely to resolve spontaneously with time. Thus short-term relief, as noted by the authors of the AHRQ report, would be very appropriate and expected for a disc herniation. To evaluate the long-term effects in this population would be as flawed as evaluating the long-term effectiveness of antibiotics for pneumonia, as it is likely that 6-12 months following an infection all patients are better regardless of the treatment regimen.

The work group’s Key Question #1 illustrates the deficiencies in stratification by diagnosis. This question asks “In patients with low back pain, what is the effectiveness of …” This question is based on a symptom that is representative of a variety of diagnostic etiologies. While the authors did state that they considered factors that may present a favorable outcome, they clearly included studies in their analysis that evaluated symptoms rather than diagnoses. In explaining this approach the authors note that: “In the majority (>85%) of patients with low back pain, symptoms cannot be attributed to a specific disease or spinal pathology.” Their reference for this statement was an article from 2002 that is not the original source of data for this statement, rather it was actually a synopsis of a workshop on idiopathic low back pain from 1982. Hence this position, was not supported by original research. In order to be accurate, we reviewed the original article from 1982, and the authors did note that “estimates of the proportion of all low-back pain that has no definite etiology range widely from about 20% to 85%”. It appears that this is a misquote of a 30 year-old opinion piece. In the same vein, relative to use of inappropriate literature, the assessment appears to rely on a manuscript that predates both modern MRI scanning and the current use of image-guided diagnostic injections, both of which have been repeatedly shown to assist in the diagnosis of spine pathology. Literature from 20-30 years ago was used that merely evaluated a symptom-based population with non-specific techniques including blind injections. While this literature was appropriate and cutting edge at the time of publication, it is not reflective of modern medicine and has introduced flaws in assessment of these procedures.

**Advances in Procedural Technology**

One other concern raised has been the inappropriate grouping of procedural techniques -- specifically the use of image guidance. The reliable placement of steroids into the epidural space requires image guidance. The failure rate of “blind” (non-image guided) needle placement has been studied by several authors. Even in experienced hands, injection of contrast after blind needle placement, demonstrated needle placement during epidural injections was incorrect 25% of the time. Stitz determined in a study of 54 consecutive caudal injections without fluoroscopic guidance, successful injection placement on the first attempt occurred in 74.1% of the patients. Renfrew also prospectively evaluated 316 caudal approach epidural steroid injections given by staff radiologists and residents over a one-year period and noted that of 111 procedures performed by physicians who had given fewer than 10 epidural steroid injections, 53 (47.7%) resulted in correct nonfluoroscopically directed placement of the needle. For physicians who had performed between 10 and 50 such procedures, 62 (53.4%) of 116 had correct nonfluoroscopically directed placement. For staff physicians, 55 (61.7%) of 89 placements were correct. Even when the sacral hiatus was easily palpated and a staff physician was confident that he or she was within the epidural space, fluoroscopy revealed incorrect placement 14.2% of the time (seven of 49 procedures). In addition, when the needle was positioned within the sacral canal and no blood was evident on Valsalva maneuver or aspiration, the injection was venous in 29 of 316 procedures (9.2%). Price studied 200 consecutive patients referred for an epidural injection and found only 64% of
caudal epidural injections were correctly placed (p< 0.001). Obesity was associated with a reduced chance of successful placement [odds ratio (OR) 0.34 (95% confidence interval (CI) 0.17 to 0.72) BMI >30 v BMI <30]. Bartynski retrospectively studied 74 LESI procedures and found that only 55 of 74 LESI procedures (74.3%), air pressure resistance was first lost upon appropriately entering the lumbar posterior epidural space. Confirmation of tip position was made with nonionic contrast medium injection in an AP and lateral epidurogram. Manchikanti studied 100 consecutive patients and noted successful injection placement without fluoroscopic visualization was confirmed on subsequent fluoroscopic visualization in 77% of patients. However, intravenous placement of the needle was noted in 14% of the patients with positive flashback and aspiration in only half 50% of these patients. Mehta used x-ray monitoring to confirm the accuracy of extradural block in 100 patients who attended the Pain Relief Clinic for treatment of a variety of different conditions. Loss of resistance, used to identify entry into the extradural space was then confirmed with contrast injection correct needle placement was noted in only 66 of 87 (79.5%) patients. Collectively this large body of work repeatedly demonstrates that non-image guided injections are inaccurate. Given the goal of an injection is to deliver an aliquot of medication to a specific target tissue, a consideration of non-specific injections as equal to image guided injections is inappropriate in modern medicine or in any review f the literature.

Of the 29 studies included in the AHRQ report as providing evidence on efficacy of epidural steroid injections versus placebo, there were 15 interlaminar epidural steroid injections, of which only one used fluoroscopic guidance. Of the 9 caudal injection studies, only one reported fluoroscopic guidance. Of the 5 transforaminal epidural steroid injection (TFESI) studies, all utilized fluoroscopic guidance. Therefore, it is worth noting that the body of evidence cited in the AHRQ review, addressing efficacy of epidural steroid injections, involves injection of steroid into an unknown tissue space, with a high probability of never reaching the site of inflammation.

Statistical Analysis
The authors clearly state “In the primary analyses, we combined weighted mean difference (WMD) for pain and standardized mean difference (SMD) for function. The mean difference was calculated using the change between the follow-up and baseline scores.” The use of mean data mandates a normal Gaussian distribution of pain. This would not be present if a treatment resulted in a bimodal distribution of outcomes with responders and non-responders. Also normally distributed data are infrequent in these patient populations given the floor and ceiling effects of a pain scale. This is evident in two studies where mean data failed to show a difference, but the appropriate categorical data showed a difference. The use of mean data is also not in accordance with the NIH Task Force recommendation for research standards for chronic low back pain. While the authors did state they considered binary outcomes, they again only briefly mentioned this in the results and conclusions and instead focused on the invalid mean changes.

References


RESULTS

Injection Approach and Evolution of Techniques
Another consideration in the assessment of effectiveness of epidural steroid injections is the target specificity of the approach. With 3 distinct approaches included in the AHRQ review, it is important to understand that even when confirmed by image-guidance, the techniques involved in delivering steroid into the epidural space, may well have different results. It would be beneficial for the technology assessment to further explore differences in approach and the different results that can be expected from them. In addition, it is expected that over time with improvements made in technique and growing clinical expertise, pooling of evidence from 2014 with that from the 1980s may not accurately reflect the effectiveness of these procedures as they are currently performed.

Evidence of Effectiveness for Radicular Pain
When evaluating the literature on epidural steroid injections for radiculopathy and herniated disc, the authors rated 3 studies as “good.” In the case of Iverson, the study design is good, but the investigative treatment is flawed. In this case, the investigator chose three possible treatments: subcutaneous saline, epidural saline delivered via the caudal route, or epidural saline and steroid delivered via the caudal route. The authors state that these injections were performed using ultrasound guidance. While ultrasound guidance may help ensure that the needle enters the caudal space, it lacks the ability to analyze flow and ensure that the medication is reaching the desired target. It is known that the caudal epidural space is a highly vascular area and venous uptake is frequent. Successful epidural placement is known to occur in only 74-77% of patients without the use of fluoroscopy, and L5 nerve root filling with this approach is rare. Further, the decision to dilute 40mg of triamcinolone with 29 mL of saline brings into question how much steroid truly reached the target structure. Lastly while the authors used validated outcome measures, no categorical data are provided thus limiting the usefulness of the outcomes.

The Karppinen study investigates fluoroscopically-guided transforaminal epidural steroid injections compared with epidural saline. Authors have questioned the appropriateness of any epidural injection as a placebo. Despite this, the study did show early improvements with epidural steroids as compared with saline, but this author also fails to provide categorical data, which might demonstrate even more robust effects of treatment in subsets of patients.

The Cohen study of transforaminal steroids compared with transforaminal etanercept and transforaminal saline is fairly well done, though it again raises concerns about a true placebo group. In this case the author chose valid outcome measures, and provides categorical data. At one month the steroid group had better pain scores, better Oswestry Disability Index (ODI) scores, more positive categorical outcomes and substantially fewer patients requiring surgery. At 3 and 6 months the results normalized, but the steroid group continued to use less pain medication and were more satisfied with their treatment than the other groups.

Furthermore, when reviewing the ratings of the quality of the literature, there seem to be inconsistencies. For example, while the authors rated the Iverson and Karppinen papers as “good,” the Ghahreman study, which seems to meet the same criteria as these studies, was given a quality rating of “fair.” In fact, the Ghahreman study used a better technique than Iverson (fluoroscopically-guided injections) and provides categorical data on validated outcome measures.
The Friedly study was also given a “good” quality rating.\textsuperscript{13} While this study design was somewhat typical of practice patterns, the investigator included a very heterogeneous group of spine pain patients with radiographic stenosis in which the “active group” received significantly varying, non-standardized doses of steroids with various non-standardized injection techniques. Patients with buttock pain were equated with patients suffering from true radicular pain, while other possible sources of their pain (e.g., facet mediated pain or sacroiliac pain) were not properly identified and excluded in this study. Further, the investigators failed to utilize appropriate outcome measures. The measures selected were validated for back pain, they were not validated for or designed to assess the symptoms of stenosis (claudication). In addition, when reviewing the data, it becomes unclear how many of the patients in this study are being treated for leg pain vs. back pain vs. claudication. The authors also failed to provide categorical data which would allow for identification and analysis of subgroups of patients who respond better than others, as there were global improvements in pain and function with both epidural saline and epidural steroids.

\textit{Evidence of Effectiveness for Axial Back Pain}

Similar to the treatment of radicular pain, the treatment of low back pain with a targeted intervention requires an accurate patho-anatomic diagnosis. Structured reviews of the literature on this topic must take this into account when assessing the quality of the literature. Unfortunately, in this report, assessment was of the effects of facet joint injections on low back pain – a symptom, not a patho-anatomic diagnosis. Alternatively, current evidence suggests that facet joint injections are highly successful in patients with low back pain and objective radiographic evidence of a specific patho-anatomic diagnosis. For patients with radiographic evidence of either joint synovitis or a facet joint synovial cyst, prospective studies show positive outcomes\textsuperscript{14,15,16} and demonstrate half or more of these patients can avoid surgery\textsuperscript{17,18} and maintain good results at long-term follow-up.\textsuperscript{19,20} Two of these studies are prospective randomized controlled trials that do not appear in the AHRQ report.\textsuperscript{15,16}

\textbf{References:}


DISCUSSION/CONCLUSIONS

In reviewing the conclusions, NASS is concerned that they are misleading due to flaws in the assessment, as outlined below.

In relation to earlier discussions about evidence quality related to RCTs, observational studies, regardless of quality and methodology, were excluded. The quality of RCTs examined is important and not considering other well-done evidence may not allow for as objective of a review process as possible. Study design per se guarantees neither quality nor valid data; the heterogeneity of intra group characteristics, varieties in delivery of care, lack of disease specific standardization of care, and ethical considerations- to name only a few confounders- can all undermine the outcome data despite a rigorous study design. A well conducted controlled trial can yield higher levels of evidence than a small, poorly conducted or methodologically flawed RCT.

Several studies suffered from a lack of a requirement for image guidance, which could dramatically alter the technical success of the injection and therefore conclusions regarding efficacy. Others either inappropriately or inadequately defined the pathology or symptomology for which the injections were being performed. The lack of image guidance affecting the successful delivery of steroids to the anatomical target is well documented.1-7

As noted in the AHRQ’s Methods Guide for Effectiveness and Comparative Effectiveness Reviews “the interpretation of the evidence and the limits of interpretation are important. Equivalence of different treatments for a group of patients on average does not necessarily imply they are equivalent for all individuals. Attempts to explore subgroups for which benefits or harms of specific interventions vary may be needed.” Patients with radicular pain were not differentiated from those who may have had somatic leg pain from sources other than the lumbar nerve root. Without a requirement for appropriate imaging (MRI, CT) to determine if there is pathology that could involve the associated lumbar nerve root, this distinction cannot be reliably made. Several studies cited in the references did not require imaging correlation to differentiate the possible origins of lower extremity symptoms,12,16,18,37,38 didn’t specify the type of imaging used 9,10,15,20,23-26,28,29,39-41 or used an imaging modality (plain X-ray) that would not have been able to adequately evaluate disc or lateral recess architecture 8,19,42 which would be the most common sources of radicular lower extremity pain.

When attempting to determine the effectiveness of a given treatment, it is often necessary to examine beyond the mean response within comparative groups to determine if there were respondents within a given treatment population that did experience a clinically significant benefit, even when the averaged mean response appear equivalent. The trials cited in this report comparing TFESI to ILESI failed to adequately examine the subgroup populations.

Given the impact on patient care, it is imperative that practitioners and patients alike fully understand the risk and benefits of a particular treatment and other treatment options. Answering questions about the appropriateness of therapy requires consideration of risks, benefits, and costs of treatment, and again according to the tenants of evidence based medicine, must include individual patient level decision-making.1 Spinal corticosteroid injections have been shown to be very safe when done appropriately in large cohorts of over 20,000 consecutive subjects.6,7 Recent studies have also demonstrated reduced overall costs in patients that receive epidural injections for their pain, mainly attributed to a decrease in loss of productivity.8 This is in contrast to other treatment options for lumbar spine disorders. There were 14,800 opioid related deaths in the United States in 2008.9 More than
103,000 individuals are hospitalized annually in the United States for NSAID-related serious GI complications, with 16,500 NSAID-related deaths occurring each year in the United States among patients with rheumatoid arthritis and osteoarthritis. We agree that injections do not alter spine structural changes that may or may not be associated with pain and functional loss, but they do provide short term relief of symptoms and can reduce the need for surgery in patients with structural changes. However, there are limits to the number of injections over time and certain disorders (axial low back pain without an associated structural change) should not be treated with injections.

NASS thanks the AHRQ for this opportunity to comment and encourages the agency to strongly consider this feedback and the impact it has on the recommendations made.

References:


