AHRQ on the Draft Technology Assessment
“Pain Management Injection Therapies for Low Back Pain”

Comments Submitted by the International Spine Intervention Society
December 15, 2014

METHODS
We commend the authors for reviewing and synthesizing a large volume of literature. There are however major flaws in the methodology of this report that significantly limit its usefulness.

Corruption of Evidence Based Medicine (EBM) Principles
The work group appeared to take the proverbial high ground by their sole utilization of randomized controlled trials (RCTs) for determination of clinical effectiveness of injectable corticosteroids. This is unfortunately a corruption of evidence based medicine, which demands the utilization of the best available evidence, not only RCTs. This is exemplified by Sackett, who stated: "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." Therefore it is imperative that all 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 required to inform for which patient subgroups a given intervention may be effective.

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." 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” and that "ignoring the evidence from observational studies is not a viable option". 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. This work group’s decision to utilize only RCTs is unfortunate, as there are multiple, methodologically rigorous studies which
included large cohorts of consecutive subjects that offer additional insights into the clinical effectiveness of these procedures.

Any literature review that is restricted to RCTs must come with an appropriate warning. The warning should include that limiting the review to RCTs skews the results to only those findings from RCTs and does not provide a balanced view of the published literature. As a result, it should also specify that the results of the review are not sufficient to inform treatment guidelines or policy. For this reason, multispecialty societies such as the North American Spine Society and the International Spinal Intervention Society have developed treatment guidelines stemming from a full assessment of the published literature. Such guidelines are appropriately constructed to inform medical treatment decisions and health policy.

Restricting a technology assessment to only RCTs ignores many high-quality observational studies. Specific to an assessment of spinal injection therapies, many high-quality studies are excluded from this review. These studies 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 >2,000 subjects. These are just a few examples, and many other examples exist. All are ignored in 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. The populations are not simply characterized by a symptom, such as back pain or sciatica. They have a radiographically-confirmed pathoanatomic diagnosis that is responsible for their symptoms. This is a critical issue in the assessment of any study involving a targeted intervention.

In addition to the work group’s mistake in limiting this review to RCTs, it is imperative to recognize that study methodology is meaningless unless the procedures being assessed are 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 and must be considered in parallel with, or even precede, evaluation of study design in assigning value to a study.

There are also other significant concerns regarding methodology that must be taken into consideration including: underlying diagnosis and its natural history, heterogeneity of procedural techniques along with the use of imaging, and even statistical analysis.
Inadequate Assessment by Diagnosis

It is imperative to recognize that 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.\(^6,7\) Alternatively, spinal injections that treat back pain without a confirmed anatomic diagnosis yield poor results.\(^8\) The distinction here is of great importance to patients with back pain, but was not adequately accounted for by the authors of the AHRQ report who repeatedly inappropriately combined diagnostic etiologies. Of the 29 studies comparing epidural steroid injections to placebo, 22 specified radicular pain alone, six 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. Again, should we withhold all antibiotics for pneumonia given the favorable natural history, or should we state antibiotics are ineffective because all subjects were better at one year follow-up? Similarly, should we withhold pain medications from patients with fractures or after orthopedic surgery, as these conditions only result in pain and have favorable natural histories?

The work group’s Key Question #1 epitomizes the fallacy of the lack of stratification by diagnosis. This question asks “In patients with low back pain, what is the effectiveness of ...” Based on the logic presented, it is unclear why in 2014 this group chose to evaluate 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 order to justify 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, however this article is not the original source of data for this statement. The original source of this statement was actually a synopsis of a workshop on idiopathic low back pain from 1982. That article was not an original research study, and contained no original data or further references, and appears to have been an expert opinion. In that original article from 1982, 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%”. Thus in an effort to justify their approach by symptoms rather than specific diagnoses, the authors of the AHRQ report misquoted a 30 year-old opinion piece. They also relied 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. Similarly the authors utilize literature from 20-30 years ago 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. To the contrary, current literature contains studies that have replicated prevalence estimates for sources of low back pain.

Inadequate Accounting for Advances in Procedural Technology

Similar to the inappropriate lumping of underlying diagnostic etiologies is the inappropriate lumping 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 1-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 lumbar epidural steroid injection (LESI) procedures and found that in 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, consideration of non-specific injections as equal to image-guided injections is inappropriate in modern medicine or in any review of the literature.

Of the 29 studies included in the AHRQ report as providing evidence on efficacy of epidural steroid injections vs placebo, there were 15 interlaminar epidural steroid injections, of which only one used fluoroscopic guidance. Of the nine caudal injection studies, only one reported fluoroscopic guidance. Of the five 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.

Inappropriate Statistical Analysis

The authors also failed to perform an appropriate 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.
Collectively, these methodological flaws render meaningless this technology assessment’s subsequent presentation of results and conclusions. Failure to establish a diagnosis, failure to assure the use of technically sound therapeutic procedures, and failure to appropriately measure outcomes of those procedures is a recipe for disaster both in medical practice and in the interpretation of medical literature. When the technology under assessment is a medical procedure, the assessors should have a firm knowledge of the technical performance of that procedure, and the pathological processes to which it is directed in contemporary practice. This is clearly not present.

References:


RESULTS
As described above, the report’s methodology yields flawed and unsupportable results.

Injection Approach and Evolution of Techniques
An important consideration in the assessment of effectiveness of epidural steroid injections is the target specificity of the approach. The failure to adequately address image guidance has been noted. With three distinct approaches (caudal, interlaminar, transforaminal) 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. The caudal and interlaminar techniques deliver medication at some distance from the target site; spread to the ventral epidural space, at the interface of the compressive lesion and the affected nerve, can be neither controlled nor guaranteed. Although the comparative effectiveness of the transforaminal approach versus the interlaminar approach was examined in five head to head trials, the authors use only inadequate continuous data in this comparison. Examination of categorical outcomes in three of the studies favored the transforaminal approach over the interlaminar approach. In one trial the dose of corticosteroid used for the interlaminar approach was twice that of the transforaminal injections. A fifth trial compared “periradicular” injections to interlaminar injections. It is not known if the “periradicular” injections provided spread of corticosteroid to the ventral epidural space, necessary for efficacy. The flawed methodology and failure to understand the nuances of technique result in the erroneous conclusion that there is no difference in effectiveness between the interlaminar and transforaminal approaches. Rather, the categorical outcomes of controlled trial evidence support the superiority of the transforaminal approach. This is supported by the clinical effectiveness of transforaminal injections documented in large observational studies, comparative effectiveness trials, and systematic reviews.

In addition, it is expected that over time with improvements in technique, technology, and growing clinical expertise, there will be changes in outcomes for procedures, which may bear the same generic description. Pooling of evidence from 2014 with that from the 1980s may do a disservice to developing an appreciation of the effectiveness of these procedures as they are currently performed.

Corticosteroid Formulation
The authors briefly examine the two controlled trials comparing corticosteroid formulations delivered by the transforaminal route, but the important clinical context is lost in the failure to examine the totality of the evidence base. The question of the comparative effectiveness, and safety, of particulate versus non-particulate steroid formulations for transforaminal epidural injections has been a critical one for interventional pain physicians, reflected in its centrality in a Food and Drug Administration’s Safe Use Initiative. Particulate steroid formulations have been associated with rare but catastrophic spinal cord infarctions; the non-particulate steroid
dexamethasone has not. A comparative effectiveness study\textsuperscript{8} and a large Mayo Clinic observational trial of >3600 consecutive transforaminal injections with a non-inferiority analysis\textsuperscript{10} showed no difference in clinical effectiveness of particulate and non-particulate steroids in the treatment of radicular pain. The limited discussion completely misses the important clinical context.

\textbf{Quality of Evidence of Effectiveness: Radicular Pain}

When evaluating the literature on epidural steroid injections for radiculopathy and herniated disc, the authors rated three studies as “good”.\textsuperscript{11,12,13} In the case of Iverson, this study design is good, but the investigative treatment is flawed.\textsuperscript{11} 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,\textsuperscript{14,15} and L5 nerve root filling with this approach is rare.\textsuperscript{15} Further, the decision to dilute 40 mg 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.\textsuperscript{12} Authors have questioned the appropriateness of any epidural injection as a placebo.\textsuperscript{16} Despite this the study did show early improvements with epidural steroids as compared with saline. This study also fails to provide categorical data, which might have demonstrated even more robust effects of treatment in subsets of patients and, indeed, a subsequent subgroup analysis did show that transforaminal epidural steroid injections were significantly effective for patients with contained herniations.\textsuperscript{17}

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.\textsuperscript{13} In this case the authors chose valid outcome measures, and provide 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 three and six 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.

Further, when reviewing the ratings of the quality of the literature, inconsistencies exist. 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. 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.

The assessment of study quality is therefore questionable. Controlled studies of epidural steroid injections have been included that do not define the pathoanatomic process to be treated, and fail to use techniques which deliver the corticosteroid to the target tissue. Outdated controlled studies have been included. Studies with varying technique have been aggregated as “epidural steroid injections”. The methodology has prevented examination of important observational trials. These failings result in the unsupported “result” that “epidural steroid injections” provide only minimal benefit in the immediate term. There is ample evidence from controlled trials, large observational trials of prospectively collected data, and systematic reviews looking at the entire evidence base, that lumbar transforaminal epidural steroid injections provide significant pain relief and functional recovery in the immediate, short, and intermediate term.1,2,3,5,6,7,8,9,10,17,18

**Quality of Evidence of Effectiveness: Axial Back Pain**

Similar to the treatment of radicular pain, the treatment of low back pain with a targeted intervention requires an accurate pathoanatomic diagnosis. Structured reviews of the literature on this topic must take this into account when assessing the quality of the literature. Unfortunately, the authors of the AHRQ report ignored this and assessed the effects of facet joint injections on low back pain – a symptom, not a pathoanatomic 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 pathoanatomic diagnosis. For patients with radiographic evidence of either joint synovitis or a facet joint synovial cyst, prospective studies show positive outcomes20,21,22.
and demonstrate half or more of these patients can avoid surgery\textsuperscript{23,24} and maintain good results at long-term follow-up\textsuperscript{25,26}. Curiously, two of these studies are prospective randomized controlled trials that do not appear in the AHRQ report.\textsuperscript{21,22}

References:

DISCUSSION/CONCLUSIONS

An initial objection to the Conclusions is that all observational studies, regardless of quality and methodology, were excluded. This is a significant error. A well-conducted observational study can yield higher levels of evidence than a small, poorly conducted or methodologically flawed RCT.

Several of the flawed studies included in the review failed to utilize image guidance, which dramatically alters the technical success of the injection and therefore the report’s conclusions regarding efficacy. It is well-documented that image guidance dramatically improves the ability to successfully deliver steroids to the anatomical target.\(^1\) Other studies either inappropriately or inadequately defined the pathology or symptomology for which the injections were being performed. Additionally, 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.”

The authors of the AHRQ report failed to heed the wisdom of the AHRQ’s established methods, which highlight the importance of identifying and exploring subgroups of patients for which benefits and harms of spinal injections may vary. Patients with radicular pain were not differentiated from those that 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,\(^8\)\(^-\)\(^12\) didn’t specify the type of imaging used\(^9\)\(^-\)\(^25\) or used an imaging modality (plain X-ray) that would not have been able to adequately evaluate disc or lateral recess architecture\(^26\)\(^-\)\(^28\) 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 appeared equivalent. The trials cited in this report comparing TFESI to ILESI failed to do this.

Lastly, given the social implications of this poorly performed assessment and implementation of any recommendations contained within, it is imperative that practitioners and patients alike fully understand the risks 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. Spinal corticosteroid injections have been shown to be very safe when done appropriately in large cohorts of over 20,000 consecutive subjects. 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. This is in stark contrast to alternative treatment options for spine pathology. The surgery-sparing effects of epidural steroid injections have been clearly demonstrated by several studies assessing effectiveness of these injections in patients who had been selected from surgical waiting lists. This outcome represents considerable cost-savings and avoidance of the risks associated with surgery. There were 14,800 opioid related deaths in the United States in 2008. 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.

References:


