Boswell ME, Trescot AM, et al. Interventional Techniques: Evidence-based Practice Guidelines in the Management of Chronic Spinal Pain. Pain Physician 2007;10:107-111.

Comments on the methodology concerning selected spinal injections:

- The majority of the studies of epidural steroid injections cited in Table 12 and Table 14 of this guideline are included in the meta-analysis of Pinto et al 2012
- These two tables illustrate the selection and analysis methods of the guideline's approach to the interventions throughout the document
- Each table lists a particular study, followed by its characteristics (whether randomized, double blind, placebo controlled), a scoring of its methodological quality using AHRQ and Cochrane criteria, then the number of patients, then initial relief, then "long term" relief at both 3 and 6 months, then results in the short term (<6 weeks) and long term (>6 weeks) characterized as "positive" or "negative"
- In contrast to Pinto, the summary is in terms of "positive" and "negative" results without giving the effect sizes, such that a statistically nominally significant result is recorded as positive, whether the effect was clinically large or clinically small
- The lack of an effect size in the tables means that a summary of the effect sizes cannot be entered into a meta-analysis, preventing the reader from learning whether the treatment effects in the literature as a whole is clinically meaningful
- Thus, while there are few conflicts between the qualitative conclusions of the included studies between Boswell and Pinto, there are differences in the interpretability in terms of how much benefit epidural steroid injections provide over all of the studies combined
- The definitions of short and long term differ between Boswell and Pinto
    - Pinto considered effects between 2 weeks and 3 months to be short term, and considered long-term benefit to be present at 12 months or more of follow-up
    - Boswell defines long term relief as lasting between 6 weeks and 3 months, with an additional long-term time frame of 6 months
- In Table I, Boswell lists designations of levels of evidence allowing a strong evidence from a single properly designed randomized controlled trial, or from multiple "properly designed" trials of smaller size, or from multiple low quality trials
- Similarly, in Table I, Boswell allows "moderate" evidence to emerge from pseudo-randomized controlled trials and from other comparative designs including the use of historical controls
- The levels of evidence used by Boswell run the risk of assigning high levels of evidence to studies likely to be influenced by biases
- The comparisons in Pinto are all with placebo, while the Boswell study includes comparisons between interventions in which both groups received an active compound
    - Placebo comparison trials are generally not favored when there is a clearly effective treatment available for the control group

    - However, when the effectiveness of epidural steroids remains in doubt, the first inferences about their effectiveness should be based on placebo comparisons
- There are studies cited by Boswell (Riew 2000 in Table 14, ref #950) which were not included in Pinto because they do not report numerical results of pain and disability, but report treatment effects in terms of p values only
    - A p value is a statistical probability of observing a particular result in a randomly selected population sample when the null hypothesis is true for the population; it is of little interest or usefulness in estimating how effective a treatment is
- Boswell reports quality scores for the Cochrane criteria, but it is not clear which scale is being used; the Cochrane scale most commonly used in the spine literature is the Cochrane Risk of Bias scale, which has 12 points, not the 10 points in whichever undefined Cochrane scale Boswell was using
    - In addition, Pinto, who used a 10 point quality scale from the physiotherapy evidence database (the commonly used and familiar PEDro scale), not only reports the overall score but has an appendix, Appendix 3, which shows how each of the 10 criteria was rated for each included study; the reader is left to guess how Boswell et al assessed whatever criteria they were using
- For these reasons, the conclusions of Boswell et al do not seriously challenge the conclusions of Pinto et al, whose methods are more in agreement with the considerations use in the Division's rating of levels of evidence