

Health Technology Assessment

**Public Comments and Responses for
Implantable Infusion Pumps
For Chronic Noncancer Pain**

August 7, 2008

**Health Technology
Assessment Program**
P.O. Box 42712
Olympia, WA 98504-2712
<http://www.hta.hca.wa.gov>

CONTENTS

<i>PEER REVIEW COMMENTS AND ECRI RESPONSES</i> _____	3
Christopher S. Hollenbeak, PhD _____	3
Roger Chou, M.D. _____	9
Dennis C. Turk, M.D. _____	17
John D. Loeser, M.D. _____	22
<i>PUBLIC COMMENTS AND ECRI RESPONSES</i> _____	25
Delfini Group _____	25

PEER REVIEW COMMENTS AND ECRI RESPONSES

July 31, 2008

Christopher S. Hollenbeak, PhD

Thank you for your careful and thorough review of our Technology Assessment entitled *Implanted Infusion Pumps for Chronic Noncancer Pain*. Please find below a point-by-point response to your comments, and explanations of how we addressed them in our report.

Reviewer Comment: This review summarizes the evidence for effectiveness and cost-effectiveness of implanted infusion pumps in patients with chronic noncancer pain. This is an excellent review, with methods for literature search and summarization documented in detail. The quality of this report is indicative of the kind of work I consistently see coming from ECRI Institute. Thanks for the chance to review and comment on it. Below I will list my recommended changes. They are numbered in order and the numbers correspond to notes in the text.

ECRI Institute Response: Thanks!

Specific Comments

Reviewer Comment: 2. Efficacy and effectiveness are frequently lumped together as in this sentence. Since these are not randomized controlled trials, it could be argued that all of the evidence presented is for effectiveness only. I would just use “effectiveness”. At the very least, say “efficacy and effectiveness” instead of “efficacy/effectiveness” since they are different.

ECRI Institute Response: We searched for controlled and uncontrolled trials, so in reference to this stage of the report we will refer to effectiveness and efficacy studies. Thereafter, since only effectiveness studies were identified, we have removed the term ‘efficacy.’

Reviewer Comment: 4. “Stable” is undefined here. You provide more details later in the body of the paper, but also in an appendix. I don’t think you should define it here because this is the executive summary, but perhaps you could insert a note so that the reader can look for the definition.

ECRI Institute Response: Several of the decisionmakers using our reports for the program only read the Executive Summary. Consequently, we have added some methodology information to the Executive Summary.

Reviewer Comment: 5. I would not try to soften this statement by saying it is something “we realize”. I would recommend just making the assertion (i.e., although the small sample size of this evidence base limits...).

Reviewer Comment: 12. See comment [5] above.

Reviewer Comment: 41. See comment [5] above

Reviewer Comment: 46. See comment [5] above.

ECRI Institute Response: Thank you for the suggestion; changes have been made.

Reviewer Comment: 6. I don’t like this statement that the confidence interval provides a better point estimate than the point estimate. First, if there is a better point estimate than the one presented, why present this one? Second, the 95% CI doesn’t really tell you where the point estimate lies. The true parameter lies in that range or it doesn’t, and we don’t know which. That’s all we can say about it. You’ve used a random effect model, which suggests that your underlying statistical model points to a super-distribution that each of the individual study distributions are drawn from. And the superdistribution is of primary interest, not the individual study distributions. The point estimate from the random effects model gives you the best point estimate for that superdistribution. I would recommend removing this language.

ECRI Institute Response: We agree with you. We have reworded the sentence to be more clear about the different purposes of point estimates and confidence intervals.

Reviewer Comment: 10. This may be a moot point, but when in-text citations are not superscripted they should be placed inside the punctuation (e.g. baseline pain scores (2)). If they are superscripted then they go outside the punctuation (e.g. baseline pain scores.²). I realize that you may be formatting them differently in the final report, but wanted to point this out and state a preference for this formatting just in case.

ECRI Institute Response: Thank you for catching this. Our usual eagle-eye proofreader wasn’t in the office that week!

Reviewer Comment: 8. The term “Function” here is somewhat vague. Do you mean functional status? Would it be better to say functional status instead?

Reviewer Comment: 47. See comment [8] above.

ECRI Institute Response: We have changed “function” to “functional status.”

Reviewer Comment: 13. I’m not sure calling this study the “1997 U.S. CEA” adequately distinguishes it from the other cost studies. Can we refer just to the author and year? If you stick with “1997 U.S. CEA” then I would recommend spelling out cost-effectiveness analysis before abbreviating with CEA.

ECRI Institute Response: We now refer to the author name and reference number.

Reviewer Comment: 17. The graphs appear to have come out a little muddled. It may just be my printout (using a mac). If it looks fine in Windows, please disregard.

ECRI Institute Response: It looked fine in Windows. No offense to your Mac intended.

Reviewer Comment: 24. Consider saying “...in the evidence base of this report”.

ECRI Institute Response: We changed the language to, “in the synthesis portion of this report,” or “in Results Synthesis,” throughout the text.

Reviewer Comment: 25. Same as 24. It just sounds funny to me as it is because it implies portions of the report are evidence-based and portions are not.

Reviewer Comment: 35. I think you really mean cohort studies here. A time series analysis is something quite different. It would look at something like measures of GDP over time.

ECRI Institute Response: We have changed the terminology to “case series” throughout the report.

Reviewer Comment: 38. A few editorial recommendations here. I think “gender distribution” sounds better than “percentage of women enrolled”. Also, this can be made two sentences. Add “..who **were** enrolled...”

ECRI Institute Response: to take up as little space as possible in the response cells, we kept the header as ‘percentage of women enrolled.’ For the second comment, changes have been made accordingly.

Reviewer Comment: 45. It's curious that this statement is not definitive. Is the estimate unstable or is it not?

ECRI Institute Response: It is an unstable estimate. Changes have been made accordingly.

Reviewer Comment: 55. Since this is a cost-effectiveness analysis I think it is critical to mention the cost perspective. Did the study use a societal perspective? Payer? I know Greg de Lissovoy and most of his studies take those perspectives.

ECRI Institute Response: It used a societal perspective; this has been added.

Reviewer Comment: 56. Because this is a cost-effectiveness study, the primary outcome is likely an incremental cost-effectiveness ratio (ICER). This should be mentioned and perhaps defined since readers of the report may not necessarily know what in ICER is.

ECRI Institute Response: The study did not report that ratio, instead it reported the cost per year of pain relief.

ECRI Institute: We appreciate the many editorial comments you provided—these matter too. We have accepted all of these changes:

Reviewer Comment: 1. This language is awkward. Maybe say “to the site for drug delivery”.

Reviewer Comment: 3. In this graph you are trying to convey to what extent values on the x axis are equal to those on the y-axis. It would help make this case if the graph were square and the 45 degree line were actually 45 degrees.

Reviewer Comment: 42. See comment [3] above.

Reviewer Comment: 7. I recommend saying “estimate” instead of figure. The term “figure” is usually reserved for graphics.

Reviewer Comment: 9. Missing punctuation.

Reviewer Comment: 11. I don't think VAS has been defined yet, so please spell it out the first time in the executive summary. Also spell it out the first time in the main body of the report.

Reviewer Comment: 14. Please provide a number for this figure.

Reviewer Comment: 15. Please number this figure in order.

- Reviewer Comment:** 16. Please add a reference for this citation.
- Reviewer Comment:** 18. Please provide a reference for this paper.
- Reviewer Comment:** 19. Please add “the”.
- Reviewer Comment:** 20. Because the abbreviation only gets used twice I don’t think abbreviating gains any efficiency. Just say “intrathecal” and “epidural”.
- Reviewer Comment:** 21. Please identify the table.
- Reviewer Comment:** 22. Please say “...this finding was **deemed** unstable upon robustness analyses.”
- Reviewer Comment:** 23. Consider using an enumerated list here.
- Reviewer Comment:** 26. I think this is the first use of the MAUDE abbreviation. Please spell it out if so. If I missed the earlier reference where it is spelled out, then please disregard.
- Reviewer Comment:** 27. Patients should be plural.
- Reviewer Comment:** 28. Consider combining these paragraphs.
- Reviewer Comment:** 29. It might be clearer if this parenthetical note was made into a footnote.
- Reviewer Comment:** 30. This whole paragraph seems somehow out of place. Is it necessary?
- Reviewer Comment:** 31. Recommend adding cost-effectiveness since it is different from just the cost.
- Reviewer Comment:** 32. Recommend “...evidence **for** efficacy...”
- Reviewer Comment:** 33. See comment [8] above. Do we really mean functional status?
- Reviewer Comment:** 34. I would recommend italicizing “a priori”. Otherwise it can be confusing.
- Reviewer Comment:** 36. See comment [26] above. Since this is not the first use of the abbreviation there is no need to define it here (but certainly a need to define it earlier!).
- Reviewer Comment:** 37. Recommend saying “this report” instead of “the text”.
- Reviewer Comment:** 39. See comment [6] above.
- Reviewer Comment:** 40. Three studies are mentioned, but only one is cited. Need to cite all three.
- Reviewer Comment:** 43. Recommend saying “ensure” rather than “see”.
- Reviewer Comment:** 48. Recommend saying “this section” instead of “the following section”.
- Reviewer Comment:** 49. The phrase “quantity of dosage change” feels very awkward. Recommend “quantity of change in dose” instead.
- Reviewer Comment:** 50. Recommend saying “of patients...”.

Reviewer Comment: 51. Recommend saying “of patients...”.

Reviewer Comment: 52. Missing punctuation.

Reviewer Comment: 53. This sentence feels awkward. Recommend “...depend on the type and severity of complications.”

Reviewer Comment: 54. Recommend “This report estimated...” rather than “Authors estimated...”

Reviewer Comment: 57. See comment [20] above.

Reviewer Comment: 58. All of these papers need citations.

Reviewer Comment: 59. Rather than say “that outcome” please specify the outcome. I don’t think it will be redundant.

Reviewer Comment: 60. See comment [6] above.

Reviewer Comment: 61. See comment [2] above.

Reviewer Comment: 62. Optimal is the adjective, optimum is a noun.

August 1, 2008

Roger Chou, M.D.

Thank you for your careful and thorough review of our Technology Assessment entitled *Implanted Infusion Pumps for Chronic Noncancer Pain*. Please find below a point-by-point response to your comments, and explanations of how we addressed them in our report.

Reviewer Comment: The report would be better organized if the key questions were re-numbered to match the order they are presented in the text (currently KQ 4 is presented before KQ 3). Also, it might be easier to follow the report if the methods were described only in a separate section, rather than partly in the results section.

ECRI Institute Response: We have re-ordered the Key Questions.

Reviewer Comment: I couldn't tell why no studies of ziconotide were included. My recollection is that the main trials included mixed populations of patients with and without cancer pain, but I think at least one of the trials reported results separately for patients with noncancer pain.

ECRI Institute Response: None of the ziconotide studies met all inclusion criteria. A major criterion that excluded many studies, including ziconotide studies, was the requirement that the studies report long-term (≥ 6 months) follow-up data.

Reviewer Comment: I strongly feel that the report overstates the reliability of case series for evaluating effectiveness. The authors argue that case series are "reasonable" because patients are unlikely to get better and have failed other treatments and the response rates are fairly impressive. However, the same argument can be used for surgery for non specific LBP, spinal cord stimulation, and other invasive treatments such as injections—but we have controlled studies and RCTs for these interventions—why should pumps be held to a lower standard? I also believe that many patients who receive pump therapy have not truly failed "all" alternative therapies such as structured, intensive multidisciplinary therapy. Case series of fusion for LBP, coccygectomy, sacral fusion, facet joint steroid injection etc all showed pretty impressive results in favor of the intervention, but beneficial effects were not borne out in controlled studies (or additional case series failed to confirm benefits). The placebo response rates are not all that impressive when the methodological shortcomings of the studies are considered, and when compared to placebo response rates for studies of other invasive interventions for patients with refractory chronic pain, which often fall within the lower limit off the 95% CIs described in the report. Coupled with the small sample sizes of almost all of the

studies included in the meta-analyses, the substantial (unexplained) heterogeneity in several of the analyses, and other methodological problems, I think the report really overstates the evidence in favor of implantable infusion pumps and could be interpreted as an argument that controlled studies are not necessary--which I do not believe to be the case. A controlled study of pump vs. intensive interdisciplinary therapy could easily be designed and would be ethical, in my opinion.

ECRI Institute Response: As we decided to look at case series in the event that no controlled studies could be identified (as approved in the Work Plan), the size of the effect was not a consideration. That we searched for controlled trials preferentially speaks to the fact that we believe controlled trials are preferable to uncontrolled trials. We have revised our text to stress that pump recipients should have exhausted all available alternatives. A comparison of intrathecal analgesia delivered by implantable infusion pumps to intensive interdisciplinary therapy would be very interesting, but we believe the only type of controlled trial that would answer Washington State's question about whether the method of opioid delivery makes a difference would be a placebo-controlled trial in which a sham device is employed. The purpose of the report was not to establish comparative effectiveness, but to evaluate the evidence regarding the safety and efficacy/effectiveness of implanted infusion pumps for chronic noncancer pain. We were assigned to produce a systematic review addressing the question, "what is the evidence?" and case series provide the only available evidence.

Given that the condition is chronic, we thought that long-term follow-up was essential for generalizability, and were satisfied that a reasonable argument could be made that the individuals included in the case series were unlikely to improve spontaneously, particularly given the timeframe (>6 months of treatment) we required.

A technology assessment that compared findings from case series and RCTs in patients with chronic back pain treated with spinal cord stimulation devices (as well as three other patient populations treated with different interventions) did not find that sample sizes were associated with outcomes, nor did they find that outcomes from case series were significantly different from RCTs. (Dalziel et al. Health Technology Assessment, NHS R&D HTA Programme, Vol .9, No.2, 2005). With this evidence from the health technology assessment program that supports NICE in mind, we do not agree that small study sizes limit the validity of the review findings, nor do we agree that the results from the case series are necessarily different than the results of RCTs would have been.

Reviewer Comment: I also think it is hard for the reader to get a sense of how much weight the authors are placing on each of the cost analyses--all of which, in my opinion, have some methodological shortcomings. For example, it is difficult to glean from the report whether the authors believe that the costs in the Kumar analysis are realistic (very low follow-up costs), whether lack of ITT in the Kumar analysis is a serious flaw (I believe it is), whether the model in the hypothetical study is realistic, whether the sensitivity analyses are appropriate, etc. Also, as far as I can tell none of the cost analyses report statistical significance of differences in costs, which would make it very difficult to interpret results, especially for the studies with small sample sizes.

ECRI Institute Response: The revised report provides more critical analysis of the cost studies, and also provides the statement made by ECRI Institute that, overall “The available evidence is insufficient to determine whether the long-term costs of implantable infusion pumps are different from the long-term costs of nonpump treatment in the management of chronic noncancer pain.” For each cost analysis, we described why we deemed it inconclusive. Also, we have mentioned any reported tests of statistical significance (or lack of reporting such tests). Further, there is a new table summarizing the three long-term cost analyses, which appears in the Executive Summary as well as the main body.

Reviewer Comment: Regarding the analyses, I believe (and so does the Cochrane Collaboration, according to their handbook, and others including Peter Juni and Matthias Egger) that meta-analyses of observational studies should only be done if there is a strong rationale to do so. I think the meta-analyses are particularly problematic in this case due to the issues with case series as described above, the substantial clinical diversity in studies (in populations, indications, outcomes assessed, interventions, etc) and the unexplained heterogeneity that you subsequently found. I think that a qualitative analysis of results (perhaps describing a range of results across studies) should be emphasized, with meta-analyses considered a secondary, less reliable analysis and not reported or very much downplayed when there are issues such as clinical diversity and statistical heterogeneity.

ECRI Institute Response: We believe that rationale was strong to meta-analyze these studies (see text we prepared for the final version of the report, below). Case series provide the only available evidence, and although not ideal, we believe that in this instance the data were worth considering.

We did not dismiss unexplained heterogeneity; rather, we always tested for it, always reported I^2 , and always refrained from drawing quantitative conclusions in its presence, even if meta-analysis was performed. Recall that in no instances of unexplained heterogeneity did we draw quantitative conclusions.

We do not agree that the small size of the studies is a limiting factor, as one of the goals of meta-analysis is to increase the power of the evidence base. Also, because it was not found to be associated with outcome by Dalziel et al. (as described above).

Although there was clinical diversity, all studies were relevant to the topic: they addressed implantable infusion pumps for chronic noncancer pain. The variation in terms of primary pain condition and drugs studies should improve external validity.

Different outcomes were analyzed separately; we did not draw any generalized conclusions. Differences in instruments used to measure the same outcomes (e.g., different quality of life scales) would have been taken care of by the use of the standardized mean difference, another advantage of using meta-analysis even when no quantitative conclusion is rendered (however, there was an insufficient quantity of data

to perform meta-analysis for quality of life or function). For pain, VAS or NRS was uniformly used in these studies for the primary outcome.

Please see the text added to the report in response to concerns regarding the use of case series:

For this report we first searched for randomized controlled trials (RCTs) or other controlled trials on the long-term harms and efficacy of implantable infusion pumps. However, no well-controlled long-term trials exist, so we evaluated the best available evidence, case series. Case series are generally considered a lower level of evidence for measuring the impact of an intervention than controlled trials. The reason is typically that, without a control group, there is no empirical estimate of what the patients' outcomes would have been if they had *not* received the treatment of interest. Thus, one would ideally have a control group in every circumstance. This is absolutely essential when patients' future outcomes are highly uncertain. However, if the natural history of a disease is stable, substantive improvement would not be expected without the intervention in question. Case series may therefore still provide meaningful information regarding a technology, especially when the natural history of the disease is well-known, and no substantial placebo effect is anticipated. This is especially true if a decision regarding the technology must be made and there is either no time to wait for controlled trial results to become available, or no controlled trials are expected.

Chronic noncancer pain patients who are candidates for receiving pain medication delivered by implantable infusion pumps have a fairly stable natural history of disease, lasting as long as a decade on average in some of the case series we identified. In addition, their course of disease would not be expected to vary as dramatically as other pain patients because pump candidates have exhausted all other available interventions for pain, including surgery where appropriate, and have not had substantial reductions in pain. These individuals are therefore resistant to not only pain-reducing treatments, but also substantial placebo effects. For these reasons, we determined that case series provide acceptable data in the absence of controlled trials. We used case series in this analysis under the assumption that patients' future outcomes would be similar to their baseline outcomes if no treatment is given.

We believe that this is a tenable approach because rigorous work based upon controlled trials suggests that the placebo effects associated with the treatment of pain are not clinically important. To bring this work to the attention of readers, we have added the following text to the executive summary, synthesis section, and discussion.

It is possible that some placebo effect may account for part of the pain relief reported. A Cochrane Review evaluated the influence of placebo interventions for clinical conditions including pain, found a possible placebo effect on reduction of patient-reported pain (although the authors note that it is unclear whether this effect size is clinically important, and that it cannot be clearly distinguished from other potential sources of bias). The size of this effect was estimated at a SMD of only -0.25 (95% CI -0.35 to -0.16), which corresponded to a change in VAS of 6/100 (or 0.6/10).(4,5) This effect size is small compared to the pooled SMD of pre-post pain scores calculated in this report, suggesting

that the reported pain relief after pump implantation was not solely due to a placebo effect.

4. Hrobjartsson A, Gotzsche PC. Placebo interventions for all clinical conditions. In: The Cochrane Database of Systematic [internet]. Issue 2. Hoboken (NJ): John Wiley & Sons, Ltd.; 2004 Apr 19 [updated 2003 Dec 22]. [accessed 2005 May 13]. [Art. No: CD003974]. Available: DOI: 10.1002/14651858.CD003974.pub2.
5. Hrobjartsson A, Gotzsche PC. Is the placebo powerless? Update of a systematic review with 52 new randomized trials comparing placebo with no treatment. *J Intern Med* 2004 Aug;256(2):91-100.

And this text on our use of meta-analysis:

In any systematic review, reviewers must decide how to summarize evidence from multiple studies. If case series satisfy our criteria and provide acceptable evidence, we do sometimes perform a meta-analysis using them. The meta-analysis of case series can follow logically from (a) the consideration of case series based on sensitivity to patients' histories and possible futures and (b) the aggregation of results using standardized meta-analytic techniques. Although there was variation across studies with respect to characteristics of enrolled patients and treatment protocols, all included studies addressed the use of implantable infusion pumps for CNCP. Therefore, provided all other criteria for meta-analysis were satisfied, we used meta-analysis to analyze the data for each outcome. Meta-analysis is known to reduce the risk of random error to produce a more reliable and precise effect estimate, and to potentially produce more generalizable results because the results from a variety of clinical contexts and settings are averaged. However, we did not necessarily use meta-analysis to arrive at a single point estimate for an effect size: In many cases we refrained from such an estimate. There are other reasons for utilizing meta-analytic techniques. These reasons for using meta-analysis apply regardless of whether the studies were all randomized and blinded studies, or a mix of blinded and unblinded studies, or a mix of randomized and nonrandomized studies, or consisted solely of case series (as in this review).

- Increasing the power of an evidence base to determine the general direction of effect (i.e., an increase or decrease in pain), especially when an evidence base is comprised of many small studies that, considered in isolation, could lead to a Type II error (concluding there is no effect when there really is one).
- Reducing the risk of random error and production of a more reliable and precise effect estimate.
- Transparent methodology for drawing conclusions, or for deeming the findings too inconclusive or unstable to enable conclusions, thereby limiting the influence of subjective judgment on data aggregation.
- Provision of formal, objective framework that can be used to investigate potential reasons for different findings across studies. Using the meta-analytic techniques of meta-regression and subgroup meta-analysis, one can investigate whether differences in outcomes are potentially associated with differences in study protocols (e.g., drug administered, duration of treatment) or characteristics of patients enrolled (e.g., most common painful condition).

This could potentially enable identification of study protocols associated with better or worse outcomes and patient groups who are most or least likely to benefit from therapy with an implantable infusion pump.

- Formal, objective methods to evaluate the consistency and robustness of conclusions.
- Avoiding the pitfalls of narrative systematic reviews, such as vote count methods in which the qualitative findings of each study in the evidence base is considered side-by-side but never pooled quantitatively or considered with respect to the sample size (i.e., precision) of each study, possibly leading to erroneous results, and subjectivity in assessing relationships between outcomes and potential moderator variables, especially as the number of studies increases.(2,3) Vote counting has been recommended as a method of “last resort,” only to be performed when effect sizes and significance levels of the studies are unavailable.(3)
- May produce more generalizable results because the results from a variety of clinical contexts and settings are averaged.

Reviewer Comment: There is a detailed explanation of how the overall quality ratings were derived, but I have concerns about using a quantitative summary score to determine quality (see Juni et al) with arbitrary cutoffs. In particular, for a type of evidence of low evidence to begin with (case series), any flaw, even a relatively minor one, could invalidate the study and might not be reflected in the summary score. A serious flaw would be more likely to invalidate a study but might only be noted as a one point difference in overall quality score.

ECRI Institute Response: In the case of a serious flaw, we do not rely solely on the quality score. Rather, we discuss the study in our internal review committee and will exclude the study from the evidence base if all members are in agreement.

Reviewer Comment: It was not clear to me how the quantity of evidence or inconsistency was explicitly defined or factored when rating the literature. In some cases 2 inconsistent studies received a “moderate” rating while 7 studies with some consistency received a “low” rating.

ECRI Institute Response: The “moderate” rating you refer to represents the quality score of the individual studies. This is not the same concept as rating the strength or stability of the evidence supporting a conclusion. In the case of the outcome with two inconsistent studies, the studies did indeed achieve “moderate” individual study quality scores, but due to the inconsistency, no conclusion was drawn for the outcome and the outcome was therefore rated “inconclusive” (not moderate).

Reviewer Comment: Throughout the manuscript, there is inconsistent use of the terms for describing quality of evidence, which should be strong, moderate, weak or inconclusive according to the appendix (low often used instead of weak). Also, the report should consistently describe both the overall qualitative and quantitative conclusions regarding the evidence--in a number of places only one or the other is described in the summary sections and executive summary and tables.

ECRI Institute Response: We have corrected any unintentional inconsistency in our descriptions of the evidence. We do, however, use two different sets of terms in our conclusions. First, we rate the strength of evidence as “strong,” “moderate,” “weak,” or “insufficient,” in regard to the presence and direction of an effect. Separately, we assess the “stability” of any quantitative estimates (i.e., magnitude of effect) we provide for continuous outcomes. Our “stability” ratings, which derive from testing the assumptions that went into the analysis (i.e., a sensitivity analysis) are rated as “high,” “moderate,” or “low” stability, or “unstable.”

Reviewer Comment: See my concerns described above regarding the reliance of small case series to generate overall conclusions regarding effectiveness/efficacy. I also have some comments about how the cost analyses were analyzed. For cost analyses based on RCTs, the authors should first assess whether the study adhered to RCT principles for minimizing bias in order to decide if cost conclusions are reliable. Just as for clinical outcomes, cost outcomes will be biased by methodological shortcomings. It is not clear that the authors used this approach. For example, in my opinion the lack of ITT analysis is a serious shortcoming of the Kumar analysis. It is also important for the authors to judge whether the cost inputs are reasonable. In the Kumar analysis, the very low follow-up costs in the pump group and the very high follow-up costs in the non-pump group (e.g., compare physical therapy and hospitalization rates) seem pretty unrealistic. Finally, just as when comparing clinical outcomes, statistical analyses should be performed to determine statistical significance of any differences. I couldn't tell if there was a statistical difference in cost estimates for any of the studies, and there was no mention of probabilities for meeting certain cost thresholds. For the study based on the administrative database, just like when evaluating an RCT, the authors should consider the validity of the study when considering whether the cost data are valid. I have concerns about the design, particularly related to how it modeled some inputs and used administrative data for others, it is also unclear to me how well it controlled for potential confounders and assembled a valid (unbiased) inception cohort etc. For the hypothetical (modeling study), it is important to assess whether you think that the model reflects reality accurately. I could not tell how you assessed the model itself (or if the study provided enough information to judge the model), and the sensitivity analysis seemed fairly limited. The last issue with the cost studies, at least the ones based on RCT data, are the small sample sizes. Cost analyses based on such small samples are typically insufficient to guide decision-making because costs tend to be so skewed. Overall, I believe based on the information presented in the report that the evidence base is insufficient to make any firm judgment regarding cost effectiveness of pump therapy because of flaws in the studies (or at least I am not convinced that the studies are high

quality based on the information currently in the report), small sample sizes for the RCTs, some questionable parameters and inputs, and lack of data on statistical significance for the cost differences that are reported.

ECRI Institute Response: We did not draw any overarching conclusions, rather, we reported conclusions (or lack thereof) on an outcome-by-outcome basis. For the cost analyses, the revised report now notes various study validity aspects that were not performed, such as blinding or concealment of allocation. We drew no conclusion about the comparative costs of pump vs nonpump, and a statement about the inconclusiveness of the evidence appears in the Executive Summary as well as the main body. For each cost analysis, we described why we deemed it inconclusive. About the small sample sizes: one was the Kumar study in Canada, and the real problems with that study did not involve sample size, but the differential allocation of patients into groups. The other small study was the Anderson trial, which did not even compare pump treatment to nonpump treatment. So the issues were more fundamental than their small sample sizes.

Reviewer Comment: I had some specific comments and minor edits/typos that are marked with track changes or comments in the accompanying Word document of the draft.

ECRI Institute Response: Changes have been made accordingly, thank you.

Again, we thank you for your careful review and suggestions for improving the report. We sincerely appreciate your contributions to this report and to improving the care for patients with chronic pain through your work on clinical practice guidelines.

August 1, 2008

Dennis C. Turk, M.D.

Thank you for your careful and thorough review of our Technology Assessment entitled *Implanted Infusion Pumps for Chronic Noncancer Pain*. Please find below a point-by-point response to your comments, and explanations of how we addressed them in our report.

Reviewer Comment: You acknowledge but you might reiterate that although you began with a specific set of questions, data were not available for you to come to specific conclusions regarding some of the ones posed.

ECRI Institute Response: We have clarified the language in our summary statement, which we believe improves the interpretability of our conclusions (and in many instances, lack of conclusions).

Reviewer Comment: The criterion of 20% that you set as clinically important is low. If you are going to use this criterion you should provide a specific reference to support it. I am much more familiar with the 30% and 50% criteria. It might also be of interest to comments on the criterion of pain severity used for inclusion in the different clinical trials and then report the percentage of patients who were below this level at the end of the trial.

ECRI Institute Response: We have changed the text to more clearly communicate that our definition that minimal clinical significance was based upon a change of 20 percentage points, not 20 percent of the baseline value. We also added the relevant citations for this definition:

In our assessment of the meaning of continuous pain scores, we define the long-term minimally clinically significant change in pain scores (smallest amount of pain relief that is meaningful) as a change of 2 on a scale of 0-10 (note: this is 20 percentage *points*, not to be confused with 20%). Methodological studies have found that this quantity of pain relief is clinically important in general pain trials(54)and for individuals with chronic musculoskeletal pain(73) and chronic low back pain(74), which are the most common causes of pain of patients seeking treatment in the case series evaluated in the *Synthesis* section.

54. Farrar JT. What is clinically meaningful: outcome measures in pain clinical trials. Clin J Pain 2000 Jul;16(2 Suppl):S106-12.
73. Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. Eur J Pain 2004 Aug;8(4):283-91.
74. Hagg O, Fritzell P, Nordwall A. The clinical importance of changes in outcome scores after treatment for chronic low back pain. Eur Spine J 2003 Feb;12(1):12-20.

We did not report all patients who had any change at all from baseline as not all changes are large enough to be considered clinically meaningful, but we did report the proportions of patients that had at least a 25% reduction in pain and those with at least a 50% reduction in pain from baseline. We added text to the method section on why a minimum of 25% change from baseline was used for the dichotomous outcomes:

For the assessment of proportions of patients attaining clinically important pain relief, we used thresholds of 25% and 50% pain relief. We used the minimum reduction of 25% pain relief long-term as clinically meaningful because this number has been identified as clinically meaningful by researchers studying the use of implantable infusion pumps for patients with CNCP in light of the challenges of treating a patient population refractory to other available therapies, and patient preferences: “The rationale for using a 25% improvement was based on our clinical observation that most patients with chronic pain for whom all, reasonable, more conservative treatments have failed are satisfied with the therapy and think the expense and risks are justified if they receive long-term relief of at least 25% of their pain.”(6)

- “Long-term success was defined as 25% or greater decrease in VAS pain intensity. In this study, we have selected a lower threshold for defining success as compared to other pain relief modalities, where the convention has been to use a 50% reduction as a threshold because this patient group has been more refractory to various pain relief therapies. They have been on long-term narcotic therapy and are quite happy and satisfied with a long-term 25% pain relief.”(68)
6. Anderson VC, Burchiel KJ. A prospective study of long-term intrathecal morphine in the management of chronic nonmalignant pain. *Neurosurgery* 1999 Feb;44(2):289-300; discussion 300-1.
 68. Kumar K, Kelly M, Pirlot T. Continuous intrathecal morphine treatment for chronic pain of nonmalignant etiology: long-term benefits and efficacy. *Surg Neurol* 2001 Feb;55(2):79-86; discussion 86-8.

In general, the case series analyzed did not have a minimum baseline pain score (see table of Inclusion/Exclusion criteria in text). However, the pooled baseline pain score of the case series analyzed for continuous pain was 8.7 (SD 2.7). This number is more frequently referred to in the final version.

Reviewer Comment: Again, the use of the criterion of 20% improvement as clinically significant needs to be justified with a citation as I am more familiar with the 30% and 50% criteria. Additionally, research has demonstrated that the higher the initial level of pain severity the larger the percent improvement required for patients to judge it as meaningful. Thus, some mention of initial pain severity and the percentage who achieve a level below the entry would be a useful addition.

ECRI Institute Response: Please refer to above ECRI Institute responses .Also, we translated follow-up pain standardized mean differences to a 0-10 VAS score to improve interpretability.

Reviewer Comment: I believe that some of your conclusions are overly positive based on the available data. Although your conclusions pain reduction seem appropriate, the absence of sufficient data on improvement in function and quality of life are significant issues as the success of treatment should be some balance between pain reduction and improvement in function. The heterogeneity of the results also limits the conclusions. Although the text acknowledges this it seems not to weigh these concerns sufficiently in the conclusions. The very small sample sizes of the studies included (excluding than Reden and Anders, only 434 patients were implanted and the average number of implanted patients/study was only 31) If only the 7 studies that included pain as an outcome are considered the total number of patients is only 143 (mean 20.43/study) warrant additional comment and temper the conclusions even further. In general the methodological quality of the studies included were weak.

ECRI Institute Response: ECRI Institute's system for rating the strength and stability of evidence distinguish qualitative and quantitative conclusions, and the rating of the strength of evidence supporting qualitative conclusions from rating of the evidence supporting the stability of quantitative conclusions. For the outcome pain (reported continuously), we drew a qualitative conclusion that implantable infusion pumps were associated with a reduction in pain, and rated the strength of that conclusion as moderate (but subsequently changed the rating to weak after changing the calculation method of the internal validity instrument). Due to substantial unexplained heterogeneity, we did not draw a quantitative conclusion, and rated the output of the meta-analysis summary statistic as unstable. Therefore, we disagree that we did not weigh heterogeneity sufficiently in the conclusions, as we refrained from drawing a quantitative conclusion due to heterogeneity. However, to clarify the point that quantitative estimates rated as "unstable" are supported by insufficient evidence to generate a conclusion, we have removed the quantitative estimates from the outcome summary statements because this appears to be confusing.

We agree that the absence of data on change in quality of life and functional status is an important limitation of our current understanding of the role of this technology. We also agree that evidence would be more convincing if more patients were enrolled in the studies. One of the advantages of meta-analysis, however, is to increase the power of the evidence base for calculations.

Reviewer Comment: If one uses the 50% or greater improvement as a measure of clinical significance then it is important to acknowledge more clearly that 59% of treated patients did not achieve this degree of improvement and the entire literature based on 7 included studies had only 144 patients, and thus only 59 achieved this criterion of pain reduction, If the 25% criteria is used, then only 6 studies with a total sample of 115 patients were included. Here the number of successful patients (meet the criterion) is only

64 patients achieved this level of pain reduction. Both of these figures indicate the limited data available and the small number of patients who actually show benefits from treatment with pumps.

ECRI Institute Response: We have revised the text to draw attention to this point in the discussion section.

Reviewer Comment: The discussion of the cost-effectiveness data fails to acknowledge that some of the assumption of annual expenses for non-pump treated patients are suspect. The authors of the papers reporting on cost-savings tend to assume very high health care utilization in the years following treatment. Using the year prior to treatment as a comparison and basis for assumptions about subsequent treatment in the years that follow assumes that the large numbers of diagnostic tests and treatments will remain constant. I think this is an erroneous assumption. I would expect that patients will be considered for a pump when the patient is at his or her worst and has been received a large number of tests and treatments. The likelihood of repeats of these tests and continuation of failed treatments over repeat years is not likely.

ECRI Institute Response: The report noted that of the three long-term cost analyses, two reported fairly similar monthly costs of nonpump treatment: \$1,420 by de Lissovoy and \$920 by Kumar. The outlier estimate, which was noted as an outlier in the report, was the \$4,055 estimated by Reden and Anders. We hypothesized that this higher cost may have been due to the possible inclusion of other types of patients such as cancer patients (or to the nonrepresentative higher costs incurred in the month prior to pump implantation), and we agree that this figure is likely an overestimate. This is one reason we heavily discounted the Reden and Anders findings in our conclusions about costs.

Reviewer Comment: The conclusion about patients being occupied following treatment, as you acknowledge, is based on very weak data. Having said this, I believe you have being overly positive in how you present conclusions on this variable.

ECRI Institute Response: We rated the evidence supporting the summary statement on occupation as inconclusive and unstable and did not draw a conclusion regarding this outcome. To clarify that we did not intend to sound overly positive, we reworded the summary statement for that outcome as:

Current evidence is insufficient to determine whether implantable infusion pumps are associated with a change in employment status among patients with chronic noncancer pain.

Reviewer Comment: On page 85 the reports comments about patients who are candidates for pumps having “no acceptable alternatives.” This phrase is not clarified what are

acceptable alternatives? What has been tried and with what results to warrant pumps as the acceptable alternative? The report also concludes that “Substantial proportions of patients reported at least 25% and/or 50% reductions in pain.” However, it could just as easily be concluded that substantial proportions (more than 50%) of patients did not achieve these levels of improvement and, moreover, these are proportions of very small numbers based on sample size.

ECRI Institute Response: Acceptable alternatives depend upon the patients’ underlying condition, the preferences of the physicians and themselves, and what is affordable for that individual. The included studies typically do not list these (see Table of Inclusion/Exclusion criteria for the clinical studies), but most reported that patients had failed oral opioid therapy and many of the studies enrolled a patient population exclusively or predominantly comprised of patients with failed back surgery syndrome who are ineligible for additional surgery. A list of conservative treatments that can be tried prior to more invasive therapy (such as that with the implantable infusion pump) are listed in the background section.

In the discussion section we have changed the proportions of patients attaining clinically significant pain relief to the calculated proportions generated through the meta-analysis, although we did not draw any conclusions regarding the proportions of patients who attained these amounts of pain relief due to an unstable evidence rating due to unexplained heterogeneity.

Reviewer Comment: The format is clear and the material is presented in a concise fashion. The review is comprehensive and reasonably balanced. The assumptions are spelled out and the data on which the conclusions are based on are clearly defined.

Again, we thank you for your careful review and suggestions for improving the report. We sincerely appreciate your contributions to this report.

August 1, 2008

John D. Loeser, M.D.

Thank you for your careful and thorough review of our Technology Assessment entitled *Implanted Infusion Pumps for Chronic Noncancer Pain*. Please find below a point-by-point response to your comments, and explanations of how we addressed them in our report.

Invited External Review Comments

Reviewer Comment: EXECUTIVE SUMMARY

Page 1, last para: “Inappropriateness for surgical therapy” is a strange phrase, since putting a pump into a patient is surgical therapy. Suggest that “no options for surgical repair of underlying pathology” or something similar be used.”

ECRI Institute Response: Thank you for proposing this clearer wording. Changes have been made accordingly.

Reviewer Comment: Page 2: Patients being treated with baclofen for spasticity often have an occluded spinal canal and the results are splendid. Only opiates and Prialt require an open canal.

ECRI Institute Response: Thank you for bringing this to our attention. Changes have been made accordingly.

Reviewer Comment: Page 5, Change in Occupation Status: The word “occupied” is not standard in this setting. Suggest “employed in or out of the home,” which will be better understood by the readers.

ECRI Institute Response: Changes have been made accordingly.

Reviewer Comment: Page 6, Medications: Medication consumption is a complex issue, since it measures the behavior of two individuals, the patient and the doctor, and one does not know which is responsible for the drug changes. For example, if the implanting doc’s policy is to taper and d/c oral opiates after implantation, whose behavior are you assessing?

ECRI Institute Response: We agree that medication consumption, as well as the use of other treatments (e.g., physical therapy, injection therapy, etc.), is highly dependent upon physician preferences and study protocols. We provide a bulleted list of this and other factors to consider when interpreting change in use of other medications before the results for that outcome are presented.

Reviewer Comment: MAIN REPORT

Page 20, 1st para: It is worth commenting on the fact that not only are all of these other drugs that are used for pain “off label,” but there are a few studies of drug stability at body temperature of 3-6 months and very few studies of the stability of mixtures of drugs wither short-term or long-term. This is important because the formation of granuloma at the catheter tip appears to be related to high drug concentration and mixtures of drugs.”

ECRI Institute Response: Changes have been made accordingly.

Reviewer Comment: Page 34, Study Selection Criteria 2: There is no data to support the idea that placebo responses are short lived only. The internal mammary artery ligation study done here in the 1960s-1970s showed a placebo response that lasted over two years. There are good reasons to require a six month observation period for chronic pain above and beyond the myth or short-duration placebo response.

ECRI Institute Response: The primary reason that data at six months or longer are considered is because implantable infusion pumps are intended to be used long-term. We have removed all language regarding the duration of a placebo response.

Reviewer Comment: Page 57: Same comments as those for Page 6 above.

ECRI Institute Response: Changes have been made accordingly.

Reviewer Comment: Page 85, line 6: change “likely” to “unlikely.” Chronic pain patients rarely spontaneously cure themselves of their symptoms.

ECRI Institute Response: Thank you for identifying this typographical error. Changes have been made accordingly.

GENERAL COMMENTS

Reviewer Comment: This is a very thorough assessment of the literature on programmable intrathecal pumps and catheters. I am not aware of significant papers that you have overlooked. I think that it is worth emphasizing for L & I that the major issue in determining outcomes is the absence of data, not the presence of data that is not supportive. Hence the thrust should be for them to set things up to collect the data that is needed to make a meaningful outcomes assessment.

Is there any way of determining if adverse events might be associated with initial equipment and that new catheters and new pumps have reduced the rates of these? For example, new pump battery lasts longer than original did; this should reduce need for pump replacement.

ECRI Institute Response: Due to differences in adverse event reporting among studies and the small total number of patients in these studies, quantifying the change in adverse events rates over time is not currently possible.

Reviewer Comment: I remain curious as to what the Health Technology Clinical Committee will recommend based upon this report.

Reviewer Comment: This is a very thorough assessment of published literature and can serve as a starting point for the discussion of the utility of intrathecal pumps and catheters for the treatment of chronic pain. I cannot identify any relevant articles that have been overlooked. It does not address the important factor of patient desires and choices for pain management, nor does it emphasize the fact that most of the patients considered for pump and catheter implantation have already failed myriad treatment options. In fact, one could argue that intrathecal catheters and pumps should be offered to chronic pain patients before many years of treatment as usual.

ECRI Institute Response: Patient preferences were not an outcome of interest specified by the Washington State Health Technology Assessment Program; however, very little on patient preferences has been published anyhow. We agree that this is an important outcome and believe that it should be considered in future studies.

We have tried to emphasize that other available treatment options were exhausted for pump candidates enrolled in the analyzed case series. This is relevant to patient inclusion/exclusion criteria, our decision to consider uncontrolled trials, and in our discussion section where we consider implantable pumps in the context that they are a treatment of 'last resort.'

Again, we thank you for your careful review and suggestions for improving the report. We sincerely appreciate your contributions to this report.

PUBLIC COMMENTS AND ECRI RESPONSES

August 1, 2008

Delfini Group

Michael E. Stuart M.D. and Sheri A. Strite

Thank you for your comments on our Technology Assessment entitled *Implanted Infusion Pumps for Chronic Noncancer Pain*. Please find below a response to your comments.

Reviewer Comment: The authors of this report state, “For this report we first searched for randomized controlled trials (RCTs) on the long-term safety and efficacy of opioids. However, no well-controlled long-term trials exist, so we evaluated the best available, case series.” After excluding studies through title, abstract and full review the authors included thirteen case series. Many of their conclusions were rated as of “moderate strength,” defined as “somewhat convincing” evidence with a “small chance that new evidence will overturn or strengthen our conclusion.”

ECRI Institute Response: The statement, “many of their conclusions were rated as of “moderate strength,” is false. In the draft report, only one conclusion was rated as moderate, that pain is reduced after long-term use of an implantable infusion pump. In the final version, the strength of this conclusion was downgraded to “weak” due to a change in the calculation method of the internal validity scores of the individual studies in the evidence base.

Reviewer Comment: Except in *extremely rare* instances of all-or-none results (meaning, for conditions where morbidity or mortality is nearly 100 percent and, with the intervention, is decreased dramatically), case series should be used *only* for hypothesis-generating and not for drawing cause and effect conclusions.

ECRI Institute Response: We realize that the use of case series requires some assumptions. In this clinical context, we believe we made a reasonable assumption that in the absence of treatment, patients' pain levels would not have dramatically improved over the course of six months or more. This assumption was based on the prior history of these patients (e.g., failed back surgery syndrome), the strict indications for pump implantation, as well as the long-term follow-up in the studies. The revised report does discuss the possibility of a small placebo effect, which is considerably smaller than the

effect reported in the studies. To clarify these points, we have modified the text to read as follows:

For this report we first searched for randomized controlled trials (RCTs) or other controlled trials on the long-term harms and efficacy of implantable infusion pumps. However, no well-controlled long-term trials exist, so we evaluated the best available evidence, case series. Case series are generally considered a lower level of evidence for measuring the impact of an intervention than controlled trials. The reason is typically that, without a control group, there is no empirical estimate of what the patients' outcomes would have been if they had *not* received the treatment of interest. Thus, one would ideally have a control group in every circumstance. This is absolutely essential when patients' future outcomes are highly uncertain. However, if the natural history of a disease is stable, substantive improvement would not be expected without the intervention in question. Case series may therefore still provide meaningful information regarding a technology, especially when the natural history of the disease is well-known, and no substantial placebo effect is anticipated. This is especially true if a decision regarding the technology must be made and there is either no time to wait for controlled trial results to become available, or no controlled trials are expected.

Chronic noncancer pain patients who are candidates for receiving pain medication delivered by implantable infusion pumps have a fairly stable natural history of disease, lasting as long as a decade on average in some of the case series we identified. In addition, their course of disease would not be expected to vary as dramatically as other pain patients because pump candidates have exhausted all other available interventions for pain, including surgery where appropriate, and have not had substantial reductions in pain. These individuals are therefore resistant to not only pain-reducing treatments, but also substantial placebo effects. For these reasons, we determined that case series provide acceptable data in the absence of controlled trials. We used case series in this analysis under the assumption that patients' future outcomes would be similar to their baseline outcomes if no treatment is given.

We believe that this is a tenable approach because rigorous work based upon controlled trials suggests that the placebo effects associated with the treatment of pain are not clinically important. To bring this work to the attention of readers, we have added the following text to the executive summary, synthesis section, and discussion:

It is possible that some placebo effect may account for part of the pain relief reported. A Cochrane Review evaluated the influence of placebo interventions for clinical conditions including pain, found a possible placebo effect on reduction of patient-reported pain (although the authors note that it is unclear whether this effect size is clinically important, and that it cannot be clearly distinguished from other potential sources of bias). The size of this effect was estimated at a SMD of only -0.25 (95% CI -0.35 to -0.16), which corresponds to a change in VAS of 6/100 (or 0.6/10).^(4,5) This effect size is small compared to the pooled SMD of pre-post pain scores calculated in this report, suggesting that the reported pain relief after pump implantation was not solely due to a placebo effect.

4. Hrobjartsson A, Gotzsche PC. Placebo interventions for all clinical conditions. In: The Cochrane Database of Systematic [internet]. Issue 2. Hoboken (NJ): John Wiley & Sons, Ltd.; 2004 Apr 19 [updated 2003 Dec 22]. [accessed 2005 May 13]. [Art. No: CD003974]. Available: DOI: 10.1002/14651858.CD003974.pub2.
5. Hrobjartsson A, Gotzsche PC. Is the placebo powerless? Update of a systematic review with 52 new randomized trials comparing placebo with no treatment. J Intern Med 2004 Aug;256(2):91-100.

Other Technology Assessment Groups also consider the use of case series in certain circumstances. As reviewed by Dalziel and colleagues (2005) for the National Health Service (NHS) (Vol 9. No. 2), it is sometimes necessary to review data from case series, such as when no RCTs are available, especially when a policy decision must be made, when duration of follow-up longer than RCTs can provide is needed, and when the technology appears sufficiently promising that waiting for RCTs is unacceptable.

Reviewer Comment: People often believe that any data is better than no data; therefore, often allowing themselves to be informed by case series in the absence of valid and clinically useful RCTs. We believe that this is a problematic approach, and we discourage the use of case series (excepting in the rare instances of all-or-none results, as noted above) for informing medical decision-making.

- Case series frequently lead to erroneous conclusions because of the extremely high potential for bias, confounding and chance effects.
- Use of case series to inform medical decision-making implies that there is a scientific basis for a decision even though case series is not scientific evidence and, therefore, is misleading to decision-makers such as policy makers, clinicians and patients who may be unaware of the pitfalls of case series.
- Because of the way they are conducted, case series most frequently support use of interventions whether they actually work or not—which increases patient harms and waste and “legitimizes” the use of potentially ineffective, expensive and costly procedures, increasing the likelihood of their becoming standard practice.
- Unproven procedures becoming standard practice decreases the likelihood that valid research will be undertaken, often because it is not understood that there actually is no valid science behind the procedure. Once a procedure is viewed as proven and/or becomes a standard, it frequently is considered unethical to withhold the procedure from patients, eliminating the ability to compare the procedure to no procedure as is needed for a truly scientific approach.
- The point of using valid science to inform medical decisions is to increase predictability about cause-effect relationships and probability of outcomes. The use of case series does not increase predictability. Rather it makes it falsely appear to increase predictability, which is misleading, and it masks the reality of medical uncertainty.
- Because of the high potential for case series to mislead, we do not believe case series to be a better basis for medical decision-making than clinical judgment.

ECRI Institute Response: We considered case series to provide acceptable, albeit weak, evidence, in light of the natural history of very chronic severe noncancer pain, as described in our response above. We did not analyze data from all available case series; rather, we applied strict inclusion criteria, and only admitted studies that satisfied these criteria and were found to be of acceptable internal validity as judged by our internal

validity assessments. Some were not, however, and the conclusions in the review are based only on a subset of the case series in the literature. The strength of the evidence was invariably weak.

A technology assessment that compared findings from case series and RCTs in patients with chronic back pain treated with spinal cord stimulation devices (as well as three other patient populations treated with different interventions) did not find that outcomes from case series were significantly different from RCTs. (Dalziel et al. Health Technology Assessment, NHS R&D HTA Programme, Vol .9, No.2, 2005). With this evidence from the HTA program that supports NICE in mind, we do not agree that the results from the case series are necessarily different than the results of RCTs would have been.

We assure you that the findings of the review are not intended to be the sole basis for medical decision making, but rather represent one piece of the puzzle towards informed medical decision making.

Reviewer Comment: Case series and observational data rarely establish that a therapy “works” or “doesn’t work.” With rare exceptions, as noted above, cause and effect conclusions should not be based on observational studies or case series. Ioannidis has pointed out that even well-done observational studies have only about a 20% positive predictive value, i.e., a 20% chance of reporting correct results, whereas well-done RCTs have a positive predictive value of approximately 85% (Reference Ioannidis JPA. Why Most Published Research Findings are False. *PLoS Med* 2005; 2(8):696-701). Hormone replacement therapy (HRT) for secondary prevention of cardiac events in women was the community standard for many years because numerous well-done observational studies reported a 35% to 50% relative reduction in subsequent cardiac events in women who chose to take HRT when compared to those who did not take HRT. The healthy-user effect was not adequately dealt with through statistical adjustments in the studies. Many other examples of erroneous conclusions being drawn from observational studies, registry data and case series exist.

ECRI Institute Response: No study design is perfect, and the limitations of all study designs and how they are executed should always be considered when interpreting the study findings.

Randomizing patients to groups provides the best chance at maximizing baseline group comparability. Both RCTs and observational studies are susceptible to other biases related to blinding, attrition, funding source, ancillary treatments, generalizability and other issues of design and conduct.

While the healthy user effect has been cited as an explanation for differences in cardiac outcomes in observational vs. controlled studies of HRT, it is not relevant to this population. In contrast, we believe that patients who have failed multiple surgical procedures and simpler medical regimens are more likely to fail a new intervention if that new intervention is continued and outcomes measured over a long period of time, as in the studies in our review. An analysis of a relevant technology assessment (spinal stimulation chronic noncancer back pain) and three other topics, the NHS report by Dalziel and colleagues did not find differences in outcomes from case series and RCTs.

Reviewer Comment: The following are examples are areas where misleading outcomes were reported based on observational data and later valid experimental data demonstrated the

erroneous conclusions: pulmonary artery catheterization, routine episiotomies, brain bypass surgery, growth hormone in the intensive care unit, encainide/flecainide for extrasystoles following myocardial infarction, antioxidants for prevention of coronary events, dopamine for the treatment of acute renal failure, sodium fluoride for the treatment of vertebral fractures, lung volume reduction surgery for the treatment of severe emphysema, balloon angioplasty for the treatment of reno-vascular hypertension, plasma exchange in patients with dermatomyositis, prevention of pulmonary emboli with vena caval filters, to name just a few. There are many examples in other areas. A recent example of misleading conclusions from observational data is the stark contrast between cardiac procedure registry outcomes data and subsequent RCT data. Registries have reported improved mortality outcomes for angioplasty compared to medical therapy. However, there are selection, performance and assessment biases in these database studies. This point is well-illustrated in the Boden RCT published in 2007 (Reference: Boden WE, O'Rourke RA, Teo KK et al; COURAGE Trial Research Group. Optimal medical therapy with or without PCI for stable coronary disease. N Engl J Med. 2007 Apr 12;356(15):1503-16. Epub 2007 Mar 26. PMID 17387127.) demonstrating that when selection and observation biases were minimized by an experimental design, no survival advantage with angioplasty when compared to medical therapy was seen.

In summary, there are many examples of significant bias, confounding and chance present in non-RCT intervention studies and a high likelihood of incorrectly drawing cause and effect conclusions when case series and observational studies are utilized for medical decision-making in the absence of valid and clinically meaningful scientific information.

ECRI Institute Response: RCTs are also susceptible to biases including confounding, differential attrition and follow-up, chance, and other biases. In addition, what RCTs gain in internal validity, they may lose in external validity. Further, RCTs are typically short-term, and the participants often are not representative of the full spectrum of patients likely to be encountered in practice. Finally, a high-profile analysis of case series and RCTs published by the NHS did not find differences in outcomes between the two study types.

Again, we thank you for your comments.

August 1, 2008

John D. Loeser, M.D.

Thank you for your careful and thorough review of our Technology Assessment entitled *Implanted Infusion Pumps for Chronic Noncancer Pain*. Please find below a point-by-point response to your comments, and explanations of how we addressed them in our report.

Invited External Review Comments

Reviewer Comment: EXECUTIVE SUMMARY

Page 1, last para: “Inappropriateness for surgical therapy” is a strange phrase, since putting a pump into a patient is surgical therapy. Suggest that “no options for surgical repair of underlying pathology” or something similar be used.”

ECRI Institute Response: Thank you for proposing this clearer wording. Changes have been made accordingly.

Reviewer Comment: Page 2: Patients being treated with baclofen for spasticity often have an occluded spinal canal and the results are splendid. Only opiates and Prialt require an open canal.

ECRI Institute Response: Thank you for bringing this to our attention. Changes have been made accordingly.

Reviewer Comment: Page 5, Change in Occupation Status: The word “occupied” is not standard in this setting. Suggest “employed in or out of the home,” which will be better understood by the readers.

ECRI Institute Response: Changes have been made accordingly.

Reviewer Comment: Page 6, Medications: Medication consumption is a complex issue, since it measures the behavior of two individuals, the patient and the doctor, and one does not know which is responsible for the drug changes. For example, if the implanting doc’s policy is to taper and d/c oral opiates after implantation, whose behavior are you assessing?

ECRI Institute Response: We agree that medication consumption, as well as the use of other treatments (e.g., physical therapy, injection therapy, etc.), is highly dependent upon physician preferences and study protocols. We provide a bulleted list of this and other factors to consider when interpreting change in use of other medications before the results for that outcome are presented.

Reviewer Comment: MAIN REPORT

Page 20, 1st para: It is worth commenting on the fact that not only are all of these other drugs that are used for pain “off label,” but there are a few studies of drug stability at body temperature of 3-6 months and very few studies of the stability of mixtures of drugs wither short-term or long-term. This is important because the formation of granuloma at the catheter tip appears to be related to high drug concentration and mixtures of drugs.”

ECRI Institute Response: Changes have been made accordingly.

Reviewer Comment: Page 34, Study Selection Criteria 2: There is no data to support the idea that placebo responses are short lived only. The internal mammary artery ligation study done here in the 1960s-1970s showed a placebo response that lasted over two years. There are good reasons to require a six month observation period for chronic pain above and beyond the myth or short-duration placebo response.

ECRI Institute Response: The primary reason that data at six months or longer are considered is because implantable infusion pumps are intended to be used long-term. We have removed all language regarding the duration of a placebo response.

Reviewer Comment: Page 57: Same comments as those for Page 6 above.

ECRI Institute Response: Changes have been made accordingly.

Reviewer Comment: Page 85, line 6: change “likely” to “unlikely.” Chronic pain patients rarely spontaneously cure themselves of their symptoms.

ECRI Institute Response: Thank you for identifying this typographical error. Changes have been made accordingly.

GENERAL COMMENTS

Reviewer Comment: This is a very thorough assessment of the literature on programmable intrathecal pumps and catheters. I am not aware of significant papers that you have overlooked. I think that it is worth emphasizing for L & I that the major issue in determining outcomes is the absence of data, not the presence of data that is not supportive. Hence the thrust should be for them to set things up to collect the data that is needed to make a meaningful outcomes assessment.

Is there any way of determining if adverse events might be associated with initial equipment and that new catheters and new pumps have reduced the rates of these? For example, new pump battery lasts longer than original did; this should reduce need for pump replacement.

ECRI Institute Response: Due to differences in adverse event reporting among studies and the small total number of patients in these studies, quantifying the change in adverse events rates over time is not currently possible.

Reviewer Comment: I remain curious as to what the Health Technology Clinical Committee will recommend based upon this report.

Reviewer Comment: This is a very thorough assessment of published literature and can serve as a starting point for the discussion of the utility of intrathecal pumps and catheters for the treatment of chronic pain. I cannot identify any relevant articles that have been overlooked. It does not address the important factor of patient desires and choices for pain management, nor does it emphasize the fact that most of the patients considered for pump and catheter implantation have already failed myriad treatment options. In fact, one could argue that intrathecal catheters and pumps should be offered to chronic pain patients before many years of treatment as usual.

ECRI Institute Response: Patient preferences were not an outcome of interest specified by the Washington State Health Technology Assessment Program; however, very little on patient preferences has been published anyhow. We agree that this is an important outcome and believe that it should be considered in future studies.

We have tried to emphasize that other available treatment options were exhausted for pump candidates enrolled in the analyzed case series. This is relevant to patient inclusion/exclusion criteria, our decision to consider uncontrolled trials, and in our discussion section where we consider implantable pumps in the context that they are a treatment of 'last resort.'

Again, we thank you for your careful review and suggestions for improving the report. We sincerely appreciate your contributions to this report.