Author's response to reviews

Title: Long-term retention on treatment with lumiracoxib 100 mg once or twice daily compared with celecoxib 200 mg once daily: a randomised controlled trial in patients with osteoarthritis [NCT00145301]

Authors:

Roy Fleischmann (RFleischmann@arthdocs.com)
Hyman Tannenbaum (hyman_tannenbaum@sympatico.ca)
Neha P Patel (nehap.patel@novartis.com)
Marianne Notter (marianne.notter@novartis.com)
Peter Sallstig (peter.sallstig@novartis.com)
Jean-Yves Reginster (jyreginster@ulg.ac.be)

Version: 3 Date: 11 January 2008

Author's response to reviews: see over
Dear Chantal

RE: MS: 9797194521648266 “Long-term retention on treatment with lumiracoxib 100 mg once or twice daily compared with celecoxib 200 mg once daily: a randomised controlled trial in patients with osteoarthritis [NCT00145301]. Roy Fleischmann, Hyman Tannenbaum, Neha P Patel, Marianne Notter, Peter Sallstig and Jean-Yves Reginster”

Thank you for considering our manuscript for publication and for the comments provided. We have assessed these comments, revised the manuscript, where appropriate, and detailed our responses below. We hope that you now find the manuscript acceptable for publication.

Reviewers comments and responses
Reviewer: Andrew Moore

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)
The authors do not tell us that they have adhered (as much as possible) to CONSORT guidelines. I think they should. While the trial is well reported, the purpose of CONSORT and other members of that family is to remind writers and readers not to forget something that may be important later, or at least to tell us why it isn't there.
Response: We appreciate the Reviewer's concern and, although it is not explicitly stated that the paper adheres to CONSORT guidelines, the manuscript was prepared in line with CONSORT. A CONSORT checklist was not requested by the journal in the instructions to authors but a checklist will accompany the resubmission and a statement will be added to the acknowledgement to indicate that CONSORT was followed.

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)
None

Discretionary Revisions (which the author can choose to ignore)
1 The use of "Prior to" rather than "before" grates. It may be common usage, but it is not good English.
Response: This text has been changed appropriately.
2 Do we need three significant figures (46.9%), or would two do just as well (47%)?
Response: It is our preference to cite data to one decimal place for accuracy.
3 There are one or two other studies that help put continuation rates into perspective. For instance Scholes et al (J Rheumatol 95 22: 708-712) performed a prospective study that showed high discontinuation rates with NSAIDs in clinical practice, and Wolfe et al (J Rheumatol 2004 31: 355-358) also examined the issue. Clinical trials may be different in nature as well as result for discontinuations. At least two large meta-analyses have looked at discontinuations in the celecoxib and valdecoxib trials, and there is long-term discontinuation data from MEDAL. The authors could perhaps push themselves to do more than just report their result, and give us the wider perspective. I could understand that they may feel that is a step too far, in which case the editors may consider an editorial, perhaps.
**Response:** As requested, the discussion has been expanded (page 19) to discuss the previous discontinuation rates studied by Scholes et al 2005 and Wolfe et al 2004 and putting these rates in perspective with traditional NSAIDs.

4 The introduction did not feel right. It was a little long, and dwelt too much on the wonders of lumiracoxib. It could be shortened and would pack more punch if it were. It needs only to say that there’s a lot of OA around, it is hard to keep folk on their pain meds, and being able to achieve longer for more would be a good thing.

**Response:** The introduction has been amended in line with the reviewer’s comment (Please see page 4 & 5).

---

**Reviewer:** Richard Day

---

**Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)**

There is not enough attention given to the serious AEs – cardiovascular and hepatic and the contrast between lumiracoxib and celecoxib that emerge. The percentage of deaths in the three groups requires more discussion. The relative risks for lumiracoxib 100mg bd versus celecoxib 200mg OD is 5 and although event numbers are small and the difference not statistically significant it needs attention in Discussion. Recall the criticism following the VIGOR publication for a similar relative risk for MI. Similarly, the cardiovascular and hepatic adverse event experiences needs to be better presented – in tables would be better. Then these need to be properly discussed. It looks to be the case that risks are less with celecoxib and although the study is not big enough, what are the implications for use in the community? It is important to deal with the liver toxicity better – the drug has been withdrawn in Australia for serious hepatotoxicity. It has been rejected by FDA recently. Hepatotoxicity is a concern yet it is brushed over in this paper. There is more enzyme elevation seen with lumiracoxib than other NSAIDs except perhaps diclofenac. What are the implications for use? Liver function tests would not be done as often as in this study but the practice lead to identification of elevation of liver enzymes greater than 3 times ULN and in some cases cessation of the drug. And this is in a trial population without comorbidities.

Table 4 would be better replaced with data on the serious AEs notably cardiovascular and hepatic, indicating exactly when they occurred in relation to the study timeline and drug treatments.

There seems to be no reason to contemplate use of lumiracoxib 100mg BID yet this is not stated or discussed in the paper. This is an important practice point.

**Response:** More attention is given to SAEs (hepatic and cardiovascular) including some discussion on hepatotoxicity observed with lumiracoxib in post-marketing surveillance in Australia (Page 20 & 21). We have also clarified that the recommended dose (100 mg once daily) of lumiracoxib should not be exceeded (including the use of 100 mg bid) for chronic use in OA (Page 21). The results section has been amended to include time to event analyses for elevations in hepatic transaminases (>3xULN) (Page 16 and Table 6), Hy’s cases (Page 16) and cardiovascular (APTC) events (Page 15 and Table 6).

---

**Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)**

The sentence in the Introduction: “The patient randomisation list was produced

Using a validated system that automated the random assignment of patient numbers to randomisation numbers in a block formation in order to ensure treatment groups were
balanced within centres.” Is extremely convoluted and difficult to follow – a rewrite would be helpful for ease of understanding.

**Response:** The sentence has been re-written to help improve understanding (Page 6).

With regard to methods, for those patients with hand or spine OA, how was the SAS scale used – “Patients completed the SAS, which comprises four 11-point scales (pain, global, difficulty with stairs and difficulty with shopping)”? Is it validated for those areas of OA involvement?

**Response:** SAS was included as an exploratory variable as it was planned to assess the validity of this scale.

In methods, “The analysis was repeated for the per protocol population (all ITT patients without major protocol violations).” Isn't this the ITT population?

**Response:** The ITT population comprised all patients who were randomised and received at least one dose of study medication, whereas the per protocol population included those patients in the ITT population who were without a major protocol violation (These definitions are on page 10). Therefore, the populations differ. The number of patients in the per protocol population has been included in the results (Page 11).

In methods, what does “to adjust for multiplicity” mean?

**Response:** Repeated statistical testing can increase the probability of a false significance. These multiple test problems (“multiplicity”) arise when there are multiple comparisons in a clinical trial, such as when there are more than two treatments/doses and/or outcome measures, or repeated measures over time. The increased probability of false significance with multiple tests can be adjusted for by statistical methods and the term “adjusted for multiplicity” is a standard way to refer to this.

Was the analysis of secondary efficacy variables performed on the ITT population?

**Response:** The methods section (Page 10) states that “All efficacy evaluations were performed on the population of randomized patients who received at least one dose of study medication, the intention-to-treat (ITT) population”. In addition, Table 3 also indicated that efficacy was assessed in the ITT population.

**Reviewer:** David Scott

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

1. The introduction about osteoarthritis is overlong; this paper will be read by experts who know most of the general background and are interested in the specific issues raised by this trial.

**Response:** the introduction has been edited down. Since the manuscript could also be read by non-experts, some general background is retained.

2. The current report is rather naïve in its approach to issues around liver toxicity. Within the EU there have been major concerns on this point and in August 2007, following consultation with the European regulatory authorities, the manufacturer wrote to healthcare professionals to inform them of new prescribing restrictions on lumiracoxib. Concern was raised worldwide after rare reports of severe liver reactions; mostly relating to daily doses higher than licensed in the EU. In the present paper the authors note that “elevations in liver enzymes of more than 3 x ULN, submitted for adjudication and considered to be related to the study drug, occurred in 1.5% of patients treated with lumiracoxib 100 mg o.d. This incidence is greater than previously observed for ALT/AST elevations >3 x ULN with lumiracoxib 100 mg o.d. in long-term clinical trials (0.91%) (Novartis: data on file, Studies 2360 [core plus extension] and 2361 [core plus extension] pooled).
However, the incidence rate for lumiracoxib 100 mg o.d. is similar to that reported in the prescribing information for many NSAIDs (1% or less), such as naproxen and ibuprofen [35,36], and less than that observed with diclofenac, which has been associated with ALT and AST abnormalities in 3.2% and 1.8% of patients, respectively [37]. They also report that ALT/AST elevations >3 x ULN occurred at a higher frequency in patients treated with lumiracoxib 100 mg b.i.d. (twice the recommended dose) (2.3%) than with lumiracoxib 100 mg o.d. (1.5%) or celecoxib 200 mg o.d. (0.4%) in the results section but not in the discussion. I suspect that the general public would think this level of hepatotoxicity is a potentially major problem against the background of the letter to healthcare professionals and the concerns by regulatory bodies. I am not certain it is wise to minimise its importance. If liver function tests indicate problems with an anti-inflammatory drug why use an agent with substantially higher risks? I am not implying that there is a correct or incorrect response to this difficult question, only that a cautious response is required in the discussion. I also think it might be wise to temper the final overall conclusions of the study.

**Response:** Much of the interest about hepatotoxicity with lumiracoxib arose after the manuscript had been finalised. However, we appreciate that this topic is now of interest and have amended the manuscript to report more hepatic data discuss hepatotoxicity (Page 16 & 20).

3. The inclusion of patients from a centre that fell below the mark for Good Clinical Practice is challenging. I do not think I can reach a conclusion on what should be done. My inclination is not to publish such work at all, though I realise this is potentially incorrect as many patients have been involved and we should all be committed to open publication. In any event I suspect more discussion is required on the pros and cons of publishing. At the very least I suspect we need to know what had gone wrong; not meeting GCP standards could be due to many reasons and not all will give rise to equal concern. Although I am certain the answer will be straightforward it would be sensible to give some indication about the standards observed in the other centres and how these were judged.

**Response:** A decision to not publish this trial could be interpreted as a lack of disclosure and leave the authors and trial sponsors open to criticism. We appreciate and understand the reviewer’s thoughts on this matter and we felt that full disclosure of both the trial and the GCP violation was necessary. We consider the data to be robust as the number of patients recruited at the center in question was a small fraction of the overall study population and the sensitivity analysis revealed that the study results did not differ significantly upon removal of the center in question. We feel that stating the reasons for the violation in the manuscript would not be helpful. However, for your information, an audit of the center revealed duplication of some laboratory results.

4. The data on the 10 most common adverse events is not helpful. The frequency of common, minor problems is not of great interest; it would be more relevant to report serious adverse events.

**Response:** We appreciate the reviewer’s suggestion but we would prefer to keep the table with the 10 most frequent adverse events. Although these events may be considered “minor”, the greater frequency of these events makes them clinically relevant.

---

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)