Reviewer's report

Title: Utility of the sore throat pain model in the multiple-dose assessment of an acute analgesic: a double-blind, randomized, placebo-controlled study of flurbiprofen 8.75 mg lozenges

Version: 3
Date: 12 February 2014

Reviewer: Christian de Mey

Reviewer's report:

Previous major concern: The authors use a test procedure to describe and compare investigational treatments under methodological restrictions that improve homogeneity and that reduce the impact of possibly confounding influences. They call this approach a “model”. I take issue with the use of such denotation. Indeed, a model is a cognitive construct that relates an experimental outcome to the underlying various effect factors (their weight and variability) and their interaction (e.g. PK-model, PKPD-model, ANOVA-model, etc.). The test procedure or routine used by the investigators does not meet this concept. I think such deflational use of the “model”-concept should be avoided (« calling my twenty years old Ford Fiesta “Mercedes” does not make it run faster ...! »).

Authors’ reply: With due respect for Reviewer 2’s comments, this “test procedure” has been termed an “acute pain model” by several authorities (in the US and Europe), pain societies, and, in fact, by Reviewer 4. The use of sore throat pain as a 'model' for the evaluation of analgesics has been described in scores of publications since its original publication (Schachtel BP, Fillingim JM, Beiter DJ, Lane AC, Schwartz LA. Rating scales for analgesics in sore throat. Clin Pharmacol Ther 1984;36:151–6, which received a lead article position by the Editor of CPT because he recognized it as a new pain model). Since then, the sore throat pain model has been recognized as one of two standard “acute pain models” for general pain (the other being dental pain). We leave the decision about using the term “model” to the Editor of the journal.

Reviewer’s reply to the authors’ reply: The authors’ arguments are impressive, but not convincing. Although this may appear rather blasphemous (considering the authors’ reply), I would like to explain my concerns in more detail:

a) Sore throat intensity was scored by the patients at one and two hours after a first dose of the assigned medication when the patients were under in-clinic surveillance; subsequently, the patients were ambulatory and were allowed to use further doses of the medication as needed (one lozenge every 3–6 hours, up to five lozenges in a 24-hour period) along with acetaminophen (to be taken as needed every 4–6 hours if there was inadequate relief from the trial lozenge). Under these conditions sore throat intensity was to be scored every hour for the remainder of the first 24-hour study period (when awake) after the first dose, and at pre-treatment, and 1 and 2 hours after each dose taken over Days 2–7. A
large effect was identified over the first 24 hours (primary criterion). However, it is obvious from the further presentation of the data and the graph of the time courses of the mean pain intensities that the effect is confined to the first 3-4 hours after the first dose. This either means a) that the investigational medication is indeed ineffective beyond about 3-4 hours after a first dose (although treatment is continued) or b) that the procedure has failed to detect a possible benefit of such treatment beyond the first dose administered. It would have been very useful to the understanding of the test procedure’s validity (as a “model”) if the authors would have provided further evidence that the first explanation indeed is the most likely whilst the second may be rejected.

b) The authors also claim “Noteworthy was the demonstration of the onset of significant differentiation of flurbiprofen 8.75 mg from placebo at the first assessment time point (1 hour) on all rating scales, peak effects at 1–3 hours, and duration lasting for up to 4 hours on two of the three scales used to measure hourly effects over 6 hours”. The chosen time axis of profiling (at hourly intervals after the first dose with the first evaluation at one hour after dosing) does not provide information on the time of onset of pain relief (an essential feature of acute pain relief); indeed, it might (and ought to) have been earlier. The chosen profiling conditions (patients being ambulatory and able to use further doses of the trial medication and/or escape medication already after 2 hours) is ill suited to identify the time of peak effect and/or the duration of effect.

c) In the analysis and discussion of the relevance of the observed effects, much emphasis has been put on reaching a 20% reduction in pain intensity from pretreatment baseline. As explained in my previous review: this is a very slight (“minimal”) reduction in pain intensity (see Cepeda et al., 2003). Such very low cut-off yields “nice” NNT, but their clinical relevance is highly questionable.

In summary: the proposed test procedure is ill designed to capture the essence of acute pain relief (time of onset of pain relief, in particular). If this kind of test procedure would be valid to evidence effects (or lack thereof) on repeated prn-dosing, then the authors ought to have provided endorsement for their finding that the investigational treatment failed to yield any significant benefit beyond 4 hours after a first dose (although treatment was continued up to 7 days). Additionally, the proposed cut-off of 20% (equivalent with minimal i.e. barely detectable) pain relief as useful to qualify pain relieving medications is highly questionable.

In consequence: my concerns about this test procedure being a sound “model” to test pain relieving medications are sustained.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.
Declaration of competing interests:

I declare that I have no competing interests.