Reviewer’s report

Title: Comparison of Hyaluronic Acid and Corticosteroid intra-articular injections for the treatment of Osteoarthritis of the hip controlled with intra-articular injections with Bupivacaine. Design of a double blinded prospective randomized controlled study.

Version: 3 Date: 17 June 2010

Reviewer: Damien Jolley

Reviewer’s report:

The manuscript presents a proposed protocol for a RCT planned to compare Hyaluronic acid (HA) to corticosteroids (CS) with a control group receiving Bupivacaine (B) by a similar intra-articular route for treatment of osteoarthritis (OA) of the hip.

First, the authors are to be commended for the decision to publish their protocol before the study is conducted. As they remark, this will ensure the rigour of the RCT, by having to meet published methods. The Cochrane Collaboration insists on publication of peer-reviewed protocols for all of their systematic reviews.

The manuscript describes a reasonably rigorous RCT design, but there are a number of omissions, and I have a number of potential improvements to suggest. The editors of BMC Musculoskeletal will address the language issues in the manuscript, but I have to state that these are numerous, and certainly need attention. I will not list them in this review.

The four principal issues of bias for Cochrane reviews of RCTs are: (1) generation of the randomisation sequence; (2) concealment of allocation; (3) effective blinding; and (4) risk of incomplete outcome data reporting (http://www.cochrane-handbook.org). We address these issues in this review.

1) Randomisation will be performed using www.randomizer.org. This is suitable for sequence generation, but it seems a set of printed lists will be generated and then downloaded from the site. This has implications for concealment of allocation (see below).

2) If lists are printed and available for view to the allocating nurse, then concealment is not possible. An alternative mechanism for the randomisation process could solve this – for example, using opaque envelopes, or real-time randomisation at infiltration.

3)

4) The authors state that the RCT will be double-blinded, because neither the patient nor the “post-operative” data collectors will be aware of the allocation. This is not the usual meaning of “double-blind”, however. See, for example, Shulz & Grimes:
In a double-blind trial, participants, investigators, and assessors usually all remain unaware of the intervention assignments throughout the trial. In view of the fact that three groups are kept ignorant, the terminology double blind is sometimes misleading. In medical research, however, an investigator frequently also assesses, so in this instance the terminology accurately refers to two categories.

THE LANCET • Vol 359 • February 23, 2002, p697

In this protocol, it appears that the investigators/surgeons will be aware of the assignment of each patient. Using Shulz & Grimes definition, this is not double-blinded. There are good reasons to blind the investigators/surgeons – knowledge of the allocated treatement may consciously or subconsciously bias the carer to better care.

5) The recognised standard for RCT data analysis is intention-to-treat (ITT). In a nutshell, this can be stated: “Patients are analysed in the groups to which they are randomised.” It is extremely concerning, therefore, to see this protocol’s third Exclusion Criterion: “Non-compliance to the study procedures and or non-completion of the study according to investigator’s judgement.” This criterion makes ITT analysis impossible. Patients must be included and analysed whether they follow the treatment protocol or not. Further, the exclusion criterion shows the importance of having investigators/carers blind to the allocated treatment. If the investigator can unilaterally choose to exclude any patient, there is an exceedingly high risk of selection bias.

From the above list, it is clear that this protocol needs some revision to avoid a study with high risk of bias. There are some other issues concerning the proposed methods and sample size.

I shall begin with the hypotheses, which are stated in terms of binary results, for example “There are differences in functional outcome (HHS & VAS) between the treatment groups (HA & CS) and the control group (B) in favour of the treatment groups, for the treatment of OA of the hip at 6 months follow-up”. A protocol needs to specify the magnitude of the minimum clinically important difference. If you choose to state the alternative hypothesis, rather than the null hypothesis, then it is your responsibility to nominate what you mean by “differences” (and just “significant differences” is invalid as well). This failure in the hypotheses statement has ramifications later when considering sample size.

When sample size is considered (under “Description methodology”), the authors choose to use a dimensionless standardised effect size of delta=0.4 to obtain their result. But this number has no clinical meaning at all. It is defined as the difference in observed means, divided by the standard deviation of the population (as in psychometric research) but does not communicate whether the patient will feel better or perform better. Only the ABSOLUTE difference in means can achieve this – for example, we declare that a change in mean VAS of 10mm or greater will be considered to be clinically important.
(Although some studies report “SMD” as defined in the Cochrane Handbook, this is recommended only when comparing studies which have different outcome measures. Within the one study, with identical outcome measures, the difference in means is the appropriate effect size for continuous measures.)

The protocol’s sample size calculations are correct, assuming the stringent alpha=0.01, power=0.8 and delta=0.4, but once again ignoring the need for ITT analysis, the protocol expects to obtain no follow-up data from about 10% of patients.

The statistical analysis plan is vague. A “linear regression model for repeated measures” is probably best replaced by a more flexible method such as that used by Qvistgaard et al (ref 15). The plan fails to mention the reporting of confidence intervals at all. In fact, confidence interval analysis is preferred throughout, making Bonferroni adjustments and interim analysis P-values unnecessary.

To summarise, this protocol describes an RCT which is at risk of bias and has unclear power calculations. The document’s English expression needs attention in several places. The authors are to be commended for subjecting their protocol to peer-review and ultimately, I am sure, to publication.

Major compulsory revisions:
1) Ensure ITT analysis
2) Revise sample size to reflect absolute effect sizes
3) Commit to reporting confidence intervals, rather than P-values

Minor Essential Revisions
1) Address allocation concealment
2) Revise blinding requirements

Discretionary Revisions
1) Consider rewording of hypotheses
2) Consider improved analysis methods
3) Revise stringent alpha (95% confidence intervals only are necessary).

Level of interest: An article of importance in its field

Quality of written English: Needs some language corrections before being published

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:
I declare that I have no competing interests.