Reviewer’s report

Title: Cost effectiveness of total knee arthroplasty from the health care providers’ perspective before and after introduction of an interdisciplinary clinical pathway - is investment always improvement?

Version: 1 Date: 24 February 2011

Reviewer: Pieter Van Herck

Reviewer’s report:

Dear Author, Dear Editor

It has been a pleasure to review this paper concerning cost effectiveness of clinical pathway introduction. Please find my comments below.

Summary of peer review

The topic is very relevant: do the benefits of quality improvement due to TKA pathway use sufficiently cover costs, as compared to no pathway use? Only a few existing studies addressed cost effects of TKA pathway use, but a true cost effectiveness analysis has rarely been performed (see Van Herck, Vanhaecht et al, 2010; J Eval Clin Pract, for a systematic review). You refer to similar meta-analysis findings. In addition your finding is intriguing since those previous lower grade estimations found (positive) saving effects, contrary to these findings. It will be essential to look further into these differences and possible explanations.

At a more abstract level this paper fits within the very relevant discussion concerning how to provide appropriate financial incentives for healthcare providers to participate in, organize, support… quality improvement initiatives. I’m immediately curious to read whether you’ll imply to throw away pathways (seems so), or whether you delved deeper into the components of the pathway that were primarily related to these findings and therefore need reconsideration within pathways (as a completely different message). Or will one of the messages be that the additional costs of quality improvement have to be additionally covered by society, since these do not cover themselves automatically? (no straightforward recommendations noted within the abstract)

The objective of the study is clearly stated and well defined, although the following points will have to be specified in more detail: definition of ‘provider’, hospital and/or medical specialists?, what about revenue/additional income effects, were opportunity costs considered, etc.

At first glance, reading through the paper, there are immediately a number of things I picked up that need revision. The paper is very well written from a traditional medical point of view (methods description, RCT style of reporting). This is not true for: intervention (pathway) and implementation description, selection of outcome(s) rationale. These are typically important issues for a
complex intervention such as a clinical pathway. In addition, the cut off criterion for defining cost effectiveness needs serious reconsideration or clarification.

I also noticed a strange time sequence which is likely to introduce bias into the results, if I interpret it correctly: info session as part of pathway one month before surgery/baseline WOMAC measurement one week before surgery/recall measurement three months after surgery. And your main outcome is based on the difference between baseline and recall... You see that the major pathway component of the info session preceded this. It likely has inflated the baseline for that pathway group and decreases positive pathway results instead of increasing them, as should be. This is a crucial point of evaluation of this study: Should the baseline WOMAC not have preceded the info session?

Finally, there is need for more discussion of how this provider perspective relates to the societal perspective, because in the end health care is a public good, paid for by society. From a public health point of view that seems even more important, next to providing the right provider incentives (your perspective). So, please consider also whether or not, in this case, both are aligned or misaligned; does society benefit in economic terms as opposed to the provider? (if the provider does suffer a loss). I'm not implying you should do separate data collection and analysis on this point, but you should discuss it at least in hypothetical terms, because this puts the value of your own study objective and therefore of the study as a whole in the right perspective.

Major compulsory revisions

Abstract:

1. Purpose paragraph: Clarify ‘provider’ as the hospital (no individual practitioner point of view, or any combination of both)
2. You should already specify further that this pathway consists mainly of education, grouped physio, etc. and not of LOS reduction or choice of medication, implant, lab tests or medical imaging. You state this in the discussion of the main text, but it should be part of the abstract also since this nuances your findings in a very important way.
3. You should also report the WOMAC baseline difference, again to nuance your findings (may be part of the rationale for the findings). (see also the strange time sequence remark above in summary of peer review)
4. Last statement of results: check this after considering the cut off comments concerning cost effectiveness definition. (this is true for other parts of cost effectiveness reporting also)
5. The conclusion of the abstract is phrased too strongly, certainly in light of these nuances. Also refrain from generalization towards all ‘patient related outcomes’. You can only conclude this for functional status.

Main text:

Introduction:
6. Since most of the literature has concentrated on LOS and other cost effects, I think it would be better to already briefly state as part of the objective which outcomes/costs are to be considered and why not the others. The further details can then be retained in the methods section as you have provided. You mention later on in the discussion why there is no LOS focus, etc. Move it forward so that readers know what to expect in advance. Also clarify here the choice of one clinical outcome measure, and not, for example, complication rates, pain status, discharge disposition, etc. (see overview in Van Herck et al, 2010, p. 45)

Patients and Methods:

7. Study design: Here the timing issue comes up: baseline test at one week before surgery, but crucial part of pathway one month before surgery (see above). Please explain to me why this potential bias as I hypothesized above has not influenced your results. Or include this as an important point of your discussion and conclusions.

8. Second paragraph, you only refer to a ‘process documentation’ here to explain how the pathway was designed and implemented (and a few references later on in your discussion). Please clarify this process. As you know some so-called pathways are just a one hour written flowchart in a record, that nobody uses. Here—and in the pathway description below- you could be more convincing about the quality of the pathway itself, how rigorously it was developed, how much of support all disciplines showed for its use. This is often very informative concerning what effects a pathway can deliver.

9. Clinical pathway description: Here you explain the content of the info session, including first physiotherapy education. I just mention it, because this again reinforces the point that this post pathway group will have improved its WOMAC already in advance of the baseline measurement.

10. You should also explain why this was voluntary. Not only does this introduce a selection bias, as you acknowledge later on, but if you consider this to be an important part of care, why should it be voluntary while everything else is not?

Primary endpoints and evaluation target parameter:

11. It should be made clear that pathway implementation costs (as you mention later on) are not included in these endpoint calculations. (I was also thinking of the opportunity costs of time spent by all professionals in pathway development, that they cannot spent in direct care)

12. Initially it was also not clear to me which costs were included, because you categorize them quite vaguely (2nd paragraph). This again relates to already having specified or not beforehand that radiology, etc. in fact stayed identical (and hence no cost difference is to be expected). Why don’t you focus and report more specifically one those costs that could have been different, based on the pathway content? Here you create great cost difference expectations, while you in fact know that only modest differences are to be expected based on a few of all relevant cost components (most of which you did not address in the pathway).
13. You have looked at cost endpoints, but you do not mention anything about revenue/additional income effects. I was initially thinking about financial revenue systems available in many countries that reward a shorter LOS. But I understand now its irrelevance in this case. You focused a lot on additional time/HRM resource use, but mention only one main saving (group physiotherapy). Where there no other increased efficiency effects for other disciplines? (or were those already captured beforehand as a consequence of the previous THA experience you describe in the discussion?)

14. The cut off explanation paragraph at the end of this section: I understand the 7500 euro/40% rationale as point of departure. But you do not explain how you reach the 4500 euro and 4000 euro expectations based on your calculations. It’s like you are asking the reader to trust you on your word that these are the attainable targets. Please specify in detail how you reached this conclusion. Does 4500 just resembles the true underlying costs or only the direct ones? Why do you envision a 500 euro reduction due to pathway use (calculated based on predicted HRM resource allocation?). This is crucial information since this defines what’s cost effective and what is not. And even more importantly, if you can justify these numbers, why should 100 euro per gained % WOMAC then suddenly count as a pass or fail test? Would even a one euro reduction not be an adequate input to define any cut-off? This choice introduces a more or less subjective criterion. I would rather see that you just report the X euro per gained % as they occur (as you in fact do in the abstract). This would imply deleting all frequency reports in % of individual rates above or below the threshold. And this would also imply not to use logistic regression in multivariate analysis, but a continuous outcome alternative (next to changes in the univariate testing). Please convince me why this more prudent approach is not recommendable. Did you test it? Non significant results?

No further comments on the statistical analysis and sample size consideration, which have been excellently reported.

Results:

Cohort comparison:

15. Please add the baseline figures when you report the WOMAC increases (especially when they are informative, such as in this case).

16. Watch out when reporting non significant trends: ‘the relation became worse’, ‘(but rather reduce)’. These are statements you should not make for non significant findings. Emphasis should lie with the fact that for these aspects no significant differences were found. Any further speculation based on one sample trends can just be due to chance or coincidence. Do not use it to reinforce a point. Just base it on the significant findings you do have.

Sub cohort analysis:

17. Also illustrate the baseline differences, including in Table 3.

18. Table 3 indeed points towards less room for improvement at baseline (due to
info session) and a voluntary self selection risk (could be a more active/knowledgeable group as you indicate later on, but also could be a group that is just the opposite: patients most in need for pre education could be more encouraged to participate. You just wouldn’t notice their maybe great improvement because it occurs before WOMAC baseline measurement. Are there not any data available recorded during the info session to check these things?

19. Please explain the minus signs within intraindividual change in Table 3. Does this mean deterioration? Please address this further in terms of its meaning as part of the discussion (the post group does seem to have less of these minus cases). This means less extreme negative outliers in terms of WOMAC outcome?

Multivariate analysis:

There was no significant difference detected through multivariate testing except the WOMAC baseline influence. The latter would confirm some of the potential explaining elements of why you did find that significant unexpected finding in univariate testing. You describe this appropriately later on.

20. However, as far as I know multivariate testing findings always precede univariate testing findings in terms of importance of reporting. Am I wrong in concluding that the WOMAC baseline difference erased all differences? Why do you then present your results in the abstract as if there is a difference (108 vs 118)? Should the message not be that there is no significant difference when taking WOMAC baseline into account? In my opinion this study is rich enough, and would provide so many lessons, that just acknowledging this non significant finding as the central result, would still make it a very valuable publication. (in the abstract you should at least also explicitly report the non signif multivariate finding, next to the signif univariate finding)

Discussion:

21. First paragraph: ‘was even found reduced’. Please explicitly specify ‘univariate’;

22. Why ‘partially’ explained? Wasn’t it completely explained, since all signif subgroup differences disappeared?

23. Next to a gradual regression to the mean effect, also consider the strange timing sequence influence as indicated above. In any case, the WOMAC baseline difference effect just seems to create less room for improvement for the post group, and hence a lower change afterwards. I do not see how this would correspond to a regression towards the mean effect. The latter is rather a random statistical phenomenon when taking multiple samples in any sequence. A baseline difference offers a more likely (in this case non random?) explanation. Or am I wrong?

24. In the last sentence of the first paragraph you present other studies’ findings a bit too positive to contrast your results. As you know many pathway studies did
not even find a significant clinical outcome effect (see examples in Van Herck et al, 2010). So your findings might be more in line with other findings instead of being inconsistent.

Very good description of design considerations.

Economic considerations:

Well discussed.

25. Please add a source to the up to 25000 euro estimation of future additional costs due to pathway development and maintenance. Is this a realistic estimation in long term? There will be a peak cost at pathway conception and implementation, but updates later on should be less costly. In addition, consider that the FTE costs for pathway coordinators are spread over multiple pathways and often they manage many other projects part of their time. I am just saying that the number 25000 euro not should be used here to sway people in convincing that then the cost (effectiveness) results will become highly negative. This needs future study, so present it in a more nuanced way.

26. The following paragraph (feel warned, …). You warn for an increased HRM investment, but this does not correspond to your cost saving findings of this study. The future likely effect of more extensive study cannot be taken for granted and be presented as if it was already part of your finding.

Minor essential revisions

Abstract:

27. Results: Currently the sentence about ‘unexpected effects of patient information/education’ needs an additional clarification sentence. Else this point remains obscure.

Main text:

Introduction:

28. Again specify ‘provider’ when formulating the study objective.

Patients and Methods:

Clinical pathway description:

29. Please specify in the last three paragraphs what was ‘new’ and what was also true for the pre group. I guess all of these points are new, but state it more explicit (e.g. a table could illustrate the content of both, next to each other/discretionary revision)

Discussion:

30. First paragraph: ‘did not increase patient related outcomes’. Please do not generalize, unless you have data on other outcomes.
Design considerations:
31. Reformulate ‘recent purpose’ in the last sentence.

Clinical considerations:
32. Delete ‘an’ in first sentence.
33. I do not understand why so much detail is given to LOS considerations here, if this falls completely outside of your current scope/measurement. Especially the a), b), c) aspects are not informative if related to LOS. They can be relevant as literature comparison in terms of the outcomes you did measure.

Discretionary revisions
34. I would move some of the parts of clinical considerations in the discussion to the introduction to clarify the context immediately (see my comments on introduction). This would create some space to elaborate further on some missing discussion points: fit of this perspective with societal perspective? Lessons for incentives? Conclusion about lessons for future studies (e.g. the important baseline/info session point), etc.

Minor issues not for publication:
35. I noticed typographical errors across the whole of the manuscript (about one per page). Please correct these errors as part of the revision process.

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.