Reviewer's report

Title: Unbearable suffering in end-of-life cancer patients with and without a request for euthanasia studied in primary care

Version: 1 Date: 12 September 2014

Reviewer: Kenneth Chambaere

Reviewer's report:

Major compulsory revisions

1. Introduction is thorough, too much so. It's much too broad in my mind and could be boiled down to only the essential, shortly describing 1) difficulties for physicians in assessing and responding to suffering; 2) unbearable suffering as central criterion, among others, for assisted dying in most jurisdictions where it is legal -focussing only or primarily on the NL where this study is set; 3) running through the numbers of EAS (requests), in cancer patients and primary care, to set the stage and to show that your patient subset is the group best represented in EAS statistics. I don't even know what the significance of the second last paragraph is.

2. Methods, measurement instrument, extra qualitative assessment of unbearable hardship: can you give more insight and explain in more detail how you qualified/deduced unbearable hardship from this information? Was it used to rule it out or to confirm or both? And who decided on this? Further in the analysis paragraph I saw that this qualitative assessment doesn't feature in your operationalisation of unbearable hardship? Important questions given it is central to the paper.

Related to the previous question: can you explain a bit more your rationale for equating a (very) serious symptom with unbearable hardship? It may be a bridge too far to deduce the impact of a symptom from its severity...or is it here that the qualitative assessment comes in? Some symptoms can be extremely severe yet bearable for the patient, and vice versa. There's an entire conceptual discussion about unbearable hardship, as the authors undoubtedly know. As you see, a bit more clarity could be provided around the operational definition of unbearable hardship.

3. The study has a low number of cases (n=17 & 47):
   * important to warn reader that statistical power is limited and the lack of significant differences does in this case certainly not mean that there are no differences between groups. Because of the low number of cases, not many significant factors were found to discuss, even the ones in the literature associated with requesting EAS (loss of control, etc). This underscores my comment that it is very difficult to show or disprove things with limited statistical power. Though it is mentioned in the limitations section, the authors should definitely make it clear in the discussion that the absence of significant findings in no way shows that the studied factors are irrelevant (as it could be read by less
knowledgeable readers). Even the conclusion that unbearable suffering is "not the dominating motive differentiating for either or not requesting EAS" may be presumptuous. Partly as a result of this, the discussion lacks 'punch' in general. Perhaps the authors could consider additional analyses, some of which I propose under 'discretionary revisions' (points 12-15).

* if significant differences are found with such low numbers there is likely a strong connection with the significant variable and can be highlighted.
* prevalence calculations are untrustworthy and this should be mentioned.

4. I also think the Discussion section could be streamlined a bit more contentwise. For instance, paragraph 7 (underlying message of a request) could be combined with paragraph 2 (performance in 1/3), linked as an explanation for non-performance (maybe you can do the same with paragraph 9 (physician responses)). Otherwise, it's difficult to see the direct connection with your study results. Also, you could ball all possible motives for an EAS request (all of which were not significantly different in your study) into one paragraph. Maybe also one paragraph with strengths (incl. originality of your research, what your study adds) and limitations together.

Minor essential revisions

5. Number of independent tests is high, Bonferroni correction should be applied to p-values of the symptom tests.

6. I found some typo's, language could be clearer, and grammar is off sometimes. The manuscript would certainly benefit from review from a native speaker.

7. Methods, analysis, second paragraph: "Only one rating per unbearable answer was permitted: the best fitting match." is all a bit confusing at this point. I presume you mean only one rater's score (on whether or not a symptom was unbearable) was used, ie the one best fitting with what? Please clarify.

8. Results: from Table 2 onwards you only have 60 instead of 64 patients. What happened to the 4 missing patients? Did they not complete the SOS-V? Maybe I missed it somewhere, but it should be mentioned.

Discretionary revisions

9. Abstract, background: last sentence is not necessary.

10. Abstract, methods: 'bimonthly' could mean both 2x per month and every 2 months. Better to say 'every two months'.

11. Methods, first paragraph, last sentence. Can you say something more about which clinical signs were used and who interpreted them?

12. Do the authors have information on when in the final 6 months the EAS request was posed? If so, wouldn't it be better to use the SOS-V symptom scores of the data point immediately after the request? We know symptoms and their unbearability can fluctuate over time, so why not use the data point closest to the
request? I understand that this would mean you are then not using the same data points as those for patients without a request, but it would be all the stronger if you could show that when a patient requests EAS, some symptoms are more severe than in patients at the very end of life not requesting EAS.

13. Did the authors think of studying differences in evolution of symptoms between patients with and without a request, taking into account when the requests were made? Though complicated, some techniques are available to do this. Now you are not exploiting the longitudinal character of the survey.

14. Results: so many items on the SOS-V scale; did the authors consider grouping them together according to larger, internally consistent categories (ideally via factor analysis or consensus-based thematic grouping) to reduce the number of items, especially considering there are virtually no significant differences in individual items?

15. Discussion: Your paragraph about balance of interacting motives: this is something you can check with your data, if you test for differences in combinations (or simply the number) of unbearable symptoms.

16. I think the findings could be very different in non-cancer patients. Perhaps this is worth a recommendation for future research as well? What do the authors think?

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

**Quality of written English:** Not suitable for publication unless extensively edited

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

**Declaration of competing interests:**

I declare that I have no competing interests