PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	European Public Assessment Report (EPAR) Summaries for the public - are they fit for purpose? A user testing study.
AUTHORS	Raynor, David; Bryant, David

VERSION 1 - REVIEW

REVIEWER	Katri Hämeen-Anttila, PhD Development Manager Finnish Medicines Agency FIMEA Finland
REVIEW RETURNED	23-May-2013

THE STUDY	page 4, line 16: Good practice is then applied – needs a reference. - Participants were recruited from a database of people who had volunteered to take part in the testing of health information materials. More information about this database should be provided: which organization is responsible of this database? Does it cover people all over UK? How do the people in the database represent people in the UK in general? - line 46-48: current or retired healthcare professionals were excluded – it is not clear if this was the case in each round? Could this be removed and be placed earlier where other excluding criteria is described (in lines 30-34)? - line 48/49: “use of literature” should be described here in more detail, how it was measured. Now the explanation comes in the results -section under the first table. - More information about where the interviews and the questionnaires took place. In page 5 (line 47) it is told that experienced trained interviews conducted them. How many interviewers were there altogether? In line 26, the authors refer to a team, however, only two members of the team are authors of this manuscript and there is not any information available about the other team members. Should they be mentioned at least in the acknowledgements? - page 4, lines 42-45 – the sentence referring to the current European guidance needs a reference - Give an example of an “indicative answer”. - Analysis of the quantitative data is missing and analysis of the qualitative data is somewhat superficial - page 6, lines 11-13: refer to table 1 - lines 23-24: These revisions were made using best practice in information writing and design – add a reference - lines 40-44: description of the mock-up of a webpage in round 4 has already been described in page 5. The summary of the rounds under the heading “Materials tested” is good, but unnecessary repetition should be avoided.

	The authors rely mainly on their own previously published papers and the number of references is low.
RESULTS & CONCLUSIONS	- results -section includes three tables which are not referred as tables - the results -section nicely combines the quantitative and qualitative results and includes number of quotes from the interview. The qualitative results from the interviews would be more exact if the number of participants would be added in sentences where the authors refer “some”; “most participants”; “several participants”; “others”; “other participants” – how many participants are these exactly? - page 12, line 50: it is not clear whether the number in the parenthesis refer to the number of answers or the number of a participant (i.e., n=4 or P4?) - As authors themselves refer to studies which show that patients want more balanced information, including benefit information, in the discussion -part – why the revised EPAR summaries did not include a heading of the benefits of Bondronat?
GENERAL COMMENTS	It covers an important topic as developing medicines information should always take into account the perspective of the target population in order to be useful and effective. The manuscript is well-written and clear. However, in its current form the manuscript seems to be more of a report of the user testing conducted rather than a scientific paper and needs to be revised. There is a lot of material that could be included as supplements for interested readers. The introduction is a description of the context of the study leading to unnecessary repetition of some issues in the later parts of the manuscript. A theoretical background is missing. This is also reflected in the discussion where eight out of eleven paragraphs do not include any references. The authors rely mainly on their own previously published papers and the number of references is low. Here follows my specific comments: Article Summary - In the summary, it is not clear which best practice in information writing and design is used (best practice according to who?). Needs a reference. Introduction - As stated above, the introduction lacks theoretical background. This could be, e.g., some theory related to patient empowerment. - The part of the introduction which describes user testing seems to suit better in the methods -section where, in fact, some of the information is repeated (e.g., page 4, lines 18-21) Methods - page 4, line 16: Good practice is then applied – needs a reference. - Participants were recruited from a database of people who had volunteered to take part in the testing of health information materials. More information about this database should be provided: which organization is responsible of this database? Does it cover people all over UK? How do the people in the database represent people in the UK in general? - line 46-48: current or retired healthcare professionals were excluded – it is not clear if this was the case in each round? Could this be removed and be placed earlier where other excluding criteria

	is described (in lines 30-34)? - line 48/49: “use of literature” should be described here in more detail, how it was measured. Now the explanation comes in the results -section under the first table. - More information about where the interviews and the questionnaires took place. In page 5 (line 47) it is told that experienced trained interviews conducted them. How many interviewers were there altogether? In line 26, the authors refer to a team, however, only two members of the team are authors of this manuscript and there is not any information available about the other team members. Should they be mentioned at least in the acknowledgements? - page 4, lines 42-45 – the sentence referring to the current European guidance needs a reference - Give an example of an “indicative answer”. - Analysis of the quantitative data is missing and analysis of the qualitative data is somewhat superficial - page 6, lines 11-13: refer to table 1 - lines 23-24: These revisions were made using best practice in information writing and design – add a reference - lines 40-44: description of the mock-up of a webpage in round 4 has already been described in page 5. The summary of the rounds under the heading “Materials tested” is good, but unnecessary repetition should be avoided. Results - results -section includes three tables which are not referred as tables - the results -section nicely combines the quantitative and qualitative results and includes number of quotes from the interview. The qualitative results from the interviews would be more exact if the number of participants would be added in sentences where the authors refer “some”; “most participants”; “several participants”; “others”; “other participants” – how many participants are these exactly? - page 12, line 50: it is not clear whether the number in the parenthesis refer to the number of answers or the number of a participant (i.e., n=4 or P4?) - As authors themselves refer to studies which show that patients want more balanced information, including benefit information, in the discussion -part – why the revised EPAR summaries did not include a heading of the benefits of Bondronat? Discussion - Discussion includes repetition of the results (e.g., page 13, lines 9-10 could be deleted), especially in the first three paragraphs. - Discussion –part lacks discussion against a theoretical background. Even though there is a lack of research evaluating the readability of the EPARs in specific, the results could be discussed in the light of previous evidence on literature evaluating other medicines / health information materials. - Limitations should be discussed more comprehensively. - Conclusions are better written in the abstract as they should answer the research questions

REVIEWER	Corrado Barbui, MD Associate Professor of Psychiatry University of Verona, Italy

REVIEW RETURNED	05-Jun-2013

THE STUDY	My opinion is that the research question is not clearly defined. In the title it is clear that the aim is to check if the EPARs are fit for purpose. In the last sentence of the introduction this general aim is further specified into two aspects: finding the key messages and understanding the key messages. In the methods it is described that two pharmacists identified 19 points considered to be the most critical information, and these points were subsequently used to check (a) finding and (b) understanding. My main concern is that this procedure assumes by definition that EPARs are fit for purpose if these two aspects are met. I think this assumption is at least debatable, as I would add a third aspect, which is if the EPARs report all the information about medicines that the public wants to access. So I would suggest three aspects: is the information that the public wants reported in the EPARs? Is it easily accessible? Is it easily understandable? As it is now, the study is based on the background assumption that what the two pharmacists identified as the most critical information actually is the most critical information for the public. I wouldn't be that sure about that. For example, an information that I personally would like to read in the EPARs is who drafted the EPAR, the EMA or the Industry? Or both? If this information is not reported, then I would personally think the EPAR is not fit for purpose. A second concern is study design. The authors selected "the public" from a database of people who had volunteered to take part in the testing of health information material. I wonder whether these individuals may be representative of "the public". Additionally, the choice of excluding individuals taking the medicine described in the EPAR, or with that condition, is not that intuitive to me, as I may think that most readers of the EPARs are actually individuals with a personal link with the condition. This issue is mentioned in the paper, but I think it should be given more consideration as it is a crucial point. A third issue is that only one EPAR was analysed. The authors reported that this document may be considered representative of most EPAR summaries, and this is important. However, I would argue that an EPAR on a medicine for cancer may be substantially different from an EPAR on a medicine for epilepsy or schizophrenia in terms of public who need to access health information, type of information needed, and readability. Therefore, what may be relevant for EPARs on a condition may not necessarily be similarly relevant for EPARs on another condition. Another point is that individuals selected as representative of "the public" could speak English to native standards. Again, I would argue that the EPARs are not only for individuals who speak English to native standards but for all European citizenships. This means that most readers may not be that familiar with the English so the findings of the present study at best applies to a minority of individuals. In the second paragraph of the methods it is reported that "good practice is then applied". It is not that clear to me the meaning of this, I would ask to add details or a reference to what it means. In the second paragraph of the introduction it is reported that the

	EPARs are developed by the EMA. Are the authors sure about this? I thought they were prepared by the Industry with input from the EMA. Can the authors provide details on the reasons why the EPAR was analysed both in print and on-screen?
RESULTS & CONCLUSIONS	In the discussion I would suggest to add a paragraph with the study limitations. I would also ask the authors to express their view in a more explicit way about whether they would recommend the approach described in the present paper to be applied by the EMA as a new standard for producing better EPARs. What about implications in terms of time and costs?
REPORTING & ETHICS	The authors reported under competing interests that they funded Luto Research which undertook the testing. Can they explain whether this is an intellectual or financial competing interest?

VERSION 1 – AUTHOR RESPONSE

Reviewer: Katri Hämeen-Anttila, PhD

1. Thank you for the opportunity to review this manuscript which tested the readability of a European Public Assessment Report (EPAR) Summary. It covers an important topic as developing medicines information should always take into account the perspective of the target population in order to be useful and effective.

- We are pleased with acknowledgement of the importance of the topic and the need to take into account the target population of medicines information.

2. The manuscript is well-written and clear. However, in its current form the manuscript seems to be more of a report of the user testing conducted rather than a scientific paper and needs to be revised. There is a lot of material that could be included as supplements for interested readers.

- We understand how the paper might read in this way – it is our intention that the changes described below will rectify this. This includes relating the study to the theory of health literacy and reference to wider related material.

3. The introduction is a description of the context of the study leading to unnecessary repetition of some issues in the later parts of the manuscript.

- We acknowledge that there is repetition in some places and have rectified that – see Point 6 below.

A theoretical background is missing. This is also reflected in the discussion where eight out of eleven paragraphs do not include any references. The authors rely mainly on their own previously published papers and the number of references is low.

- We intended this work to be pragmatic and practical in nature, but accept that reference to some theoretical background would improve the paper – see Point 5 below for more detail. Also, as a result, there is a significant increase in the number of references which support the discussion.

ARTICLE SUMMARY

4. In the summary, it is not clear which best practice in information writing and design is used (best practice according to who?). Needs a reference.

- We accept that the body of the paper needs more detail on how we define ‘best practice’ – see

below. However, we do not feel that including a reference is appropriate in the 'Article Summary'.

INTRODUCTION
5. As stated above, the introduction lacks theoretical background. This could be, e.g., some theory related to patient empowerment.

- We feel that an appropriate theoretical background relates to 'health literacy' – we have therefore introduced a new paragraph in the 'Introduction' on this topic. We have also made reference to 'health literacy' in the Discussion – see below.

6. The part of the introduction which describes user testing seems to suit better in the methods -section where, in fact, some of the information is repeated (e.g., page 4, lines 18-21)

- We agree that this is unnecessary repetition, and have removed the two sentences on page 3 (lines 38-41) which were essentially repeated on page 4.

METHODS
7. page 4, line 16: Good practice is then applied – needs a reference.

- As mentioned under Point 4 above, we accept that the paper needs to include more detail on 'good practice'/'best practice' - and which such practice was used when revising the summaries. There was originally some brief information on page 8, under 'Revising the EPAR summary'. We have now moved this earlier to page 4 in the 'Methods' – and added further detail on the sources used, and why they were chosen.

8. Participants were recruited from a database of people who had volunteered to take part in the testing of health information materials. More information about this database should be provided: which organization is responsible of this database? Does it cover people all over UK? How do the people in the database represent people in the UK in general?

- The database is that which Luto Research use to recruit members of the public to take part in user testing of health information documents. It draws from people around the City of Leeds and the wider Yorkshire region. The UK is a relatively small geographically and relatively homogeneous – there is no evidence that undertaking testing with people from different areas of the country would give difference results in user testing (provided they are chosen according to the criteria below).
- As described on page 4 (lines 34-46), participants were recruited according to criteria to ensure that there were members of both genders, across the age range for the medicine, a minority of the very educated and at least 2 who do not routinely work with written documents.
We have amended the text accordingly to give some brief further detail.

9. Line 46-48: current or retired healthcare professionals were excluded – it is not clear if this was the case in each round? Could this be removed and be placed earlier where other excluding criteria is described (in lines 30-34)?

- We agree that this is not clear - we have moved this text to earlier in this section, as suggested.

10. Line 48/49: "use of literature" should be described here in more detail, how it was measured. Now the explanation comes in the results -section under the first table.

- We have now also included the detail at this point in the text i.e. repeated the text under the first table which explains what is meant by 'use of literature'.

11. More information about where the interviews and the questionnaires took place. In page 5 (line 47) it is told that experienced trained interviews conducted them. How many interviewers were there altogether? In line 26, the authors refer to a team, however, only two members of the team are authors of this manuscript and there is not any information available about the other team members. Should they be mentioned at least in the acknowledgements?

- Further detail about where the interviews took place, and the number of interviewers has been added on page 5.
- The 'team' consisted of DKR and DB, along with SB, KB and WM – the latter three did not meet the journal's criteria for authorship, and so, with their agreement, they are included in the 'Contributors' section. We are happy to move their names to the 'Acknowledgements' section if preferred.

12. page 4, lines 42-45 – the sentence referring to the current European guidance needs a reference

- Reference added.

13. Give an example of an "indicative answer".

- An example has been added after the concept of indicative answer is introduced in the Method.

14. Analysis of the quantitative data is missing and analysis of the qualitative data is somewhat superficial

- We were mindful of the length of the paper while it was being drafted, and we feel that the level of analysis is sufficient to give the reader the appropriate level of understanding.

15. page 6, lines 11-13: refer to table 1

- Reference now made to Table 1 (subsequently re-numbered as Table 4 – see Point 18 below).

16. Lines 23-24: These revisions were made using best practice in information writing and design – add a reference

- We have now inserted the phrase '(see above)' at this point in the text – referring the reader to the added detail and explanation of good practice at the beginning of the 'Introduction' (see Point 7 above).

17. Lines 40-44: description of the mock-up of a webpage in round 4 has already been described in page 5. The summary of the rounds under the heading "Materials tested" is good, but unnecessary repetition should be avoided.

- We accept this is unnecessary repetition and have removed the second occurrence on page 6.

RESULTS
18. Results -section includes three tables which are not referred as tables

- We apologise for this omission, and have now designated each of these as tables - and consequently renumbered the existing tables.

19. The results section nicely combines the quantitative and qualitative results and includes number of quotes from the interview. The qualitative results from the interviews would be more exact if the number of participants would be added in sentences where the authors refer "some"; "most

participants"; "several participants"; "others"; "other participants" – how many participants are these exactly?

- We are pleased that this section combines well the two types of results.
- However, we do not agree that the number of participants should be added- in qualitative results such as this, our understanding of normal practice is to use such descriptive terms when presenting such results, rather than numbers.

20. Page 12, line 50: it is not clear whether the number in the parenthesis refer to the number of answers or the number of a participant (i.e., n=4 or P4?)

- We apologise for this lack of clarity – the numbers refer to the number of answers, so we have amended to read (n=4) etc.

21. As authors themselves refer to studies which show that patients want more balanced information, including benefit information, in the discussion -part – why the revised EPAR summaries did not include a heading of the benefits of Bondronat?

- This is an excellent observation which with hind-sight we would have addressed. We have now acknowledged this in the new Limitations section of the Discussion:
'We refer later in this Discussion to studies which show that patients want more balanced information, including benefit information. However, we did not include such a heading in the final version of the EPAR summary e.g. 'What are the benefits of Bondronat'? The inclusion of such a heading may need to be studied in future research on these documents.'

DISCUSSSION
22. Discussion includes repetition of the results (e.g., page 13, lines 9-10 could be deleted), especially in the first three paragraphs.

- We have removed unnecessary repetition of results in each of the first 3 paragraphs, and merged the second two paragraphs into one. Description of some of the key findings has been retained, to help frame the discussion.

23. Discussion –part lacks discussion against a theoretical background. Even though there is a lack of research evaluating the readability of the EPARs in specific, the results could be discussed in the light of previous evidence on literature evaluating other medicines / health information materials.

- As mentioned above in Point 5, we have made reference to the 'health literacy' literature in relation to this context of the study (in the 'Introduction') and again here in the Discussion, where we have added six new references, as proposed by Reviewer 1 above.

24. Limitations should be discussed more comprehensively.

- We have introduced a sub-heading 'Limitations' in the 'Discussion' and included the relevant points raised by both reviewers.

25. Conclusions are better written in the abstract as they should answer the research questions

- We have revised the final paragraph to better reflect the wording in the Abstract (although the wording in both has also been amended as a result of
Reviewer 2; Point 2 below).

Reviewer: Corrado Barbui, MD
1. My opinion is that the research question is not clearly defined.
- In the title it is clear that the aim is to check if the EPARs are fit for purpose.
- In the last sentence of the introduction this general aim is further specified into two aspects: finding the key messages and understanding the key messages.
- In the methods it is described that two pharmacists identified 19 points considered to be the most critical information, and these points were subsequently used to check (a) finding and (b) understanding.

- We agree that the research question needs clarifying and that this point closely links with Point 2 below – please see below for our response to both Points 1 and 2. The result is that we have revised the wording related to the research question, and made it consistent, in each of the 3 places noted by the Reviewer.

2. My main concern is that this procedure assumes by definition that EPARs are fit for purpose if these two aspects are met. I think this assumption is at least debatable, as I would add a third aspect, which is if the EPARs report all the information about medicines that the public wants to access. So I would suggest three aspects:
- is the information that the public wants reported in the EPARs?
- Is it easily accessible?
- Is it easily understandable?
As it is now, the study is based on the background assumption that what the two pharmacists identified as the most critical information actually is the most critical information for the public. I wouldn't be that sure about that. For example, information that I personally would like to read in the EPARs is who drafted the EPAR, the EMA or the Industry? Or both? If this information is not reported, then I would personally think the EPAR is not fit for purpose.

- This is a very fair point, and relates to the research question issue above. There is indeed a third aspect to fitness-for-purpose: does the EPAR Summary include all the information that the public wants to access. In addition, this links to the question below about whom the document should be aimed at.
- As a result, we have amended the text to remove the term 'fitness-for-purpose', and standardised on the terminology related to being able to find and understand key points of information. This means the following changes:
- TITLE: needs to change from 'EPAR Summaries for the public – are they fit for purpose?' to 'EPAR Summaries for the public – can people find and understand the information the key messages?' HOWEVER we have not changed the title at this stage, pending advice from the Editor on how to do this in the submission
- ABSTRACT Conclusions: Changed to: 'People had difficulty finding and understanding key messages in the Summary, but user testing identified the problems, and application of good practice resulted in revised formats which performed well.'
- CONCLUSIONS: Changed to: 'However, the EPAR summary document did not perform well in user testing, but the testing highlighted the problems, and application of good practice resulted in revised formats which performed well.'

3. A second concern is study design. The authors selected "the public" from a database of people who had volunteered to take part in the testing of health information material.
- I wonder whether these individuals may be representative of "the public".

- We believe the participants were representative of the public, based on the criteria through which they were selected – see response to Reviewer 1 (Point 8) above.

4. Additionally, the choice of excluding individuals taking the medicine described in the EPAR, or with that condition, is not that intuitive to me, as I may think that most readers of the EPARs are actually individuals with a personal link with the condition.
-This issue is mentioned in the paper, but I think it should be given more consideration as it is a crucial point.

- We chose to exclude individuals taking the medicine, as the stated audience by the EMA was the general public – not people taking the medicine. It is a pre-requisite of the user testing process that the target audience are the participants. However, as we indeed state in our recommendations in Box 1:
"The general public may not be the right audience. It may be more usefully aimed at people who take or are considering taking the medicine concerned and people from relevant patient groups, rather than the general public."
We are happy to include this point in the main text, but feel it is appropriately described here.

5. A third issue is that only one EPAR was analysed. The authors reported that this document may be considered representative of most EPAR summaries, and this is important. However, I would argue that an EPAR on a medicine for cancer may be substantially different from an EPAR on a medicine for epilepsy or schizophrenia in terms of public who need to access health information, type of information needed, and readability. Therefore, what may be relevant for EPARs on a condition may not necessarily be similarly relevant for EPARs on another condition.

- We accept that this is a limitation, and we have amended the existing text to acknowledge this, and moved it to the new 'Limitations' heading. We have also expanded the text, in line with the reviewer's example.

6. Another point is that individuals selected as representative of "the public" could speak English to native standards. Again, I would argue that the EPARs are not only for individuals who speak English to native standards but for all European citizenships. This means that most readers may not be that familiar with the English so the findings of the present study at best applies to a minority of individuals.

- EPAR Summaries are available in all the official languages of the EU – we have added that information to the Methods section on page 4.

7. In the second paragraph of the methods it is reported that "good practice is then applied". It is not that clear to me the meaning of this, I would ask to add details or a reference to what it means.

- See Reviewer 1; Point 7 above for our comments in relation to the similar comments of Reviewer 1.

8. In the second paragraph of the introduction it is reported that the EPARs are developed by the EMA. Are the authors sure about this? I thought they were prepared by the Industry with input from the EMA.

- Yes we are sure that EPARs and EPAR summaries are developed by the EMA – see: www.ema.europa.eu/ema/index.jsp?curl=pages/ medicines/general/general_content_000433.jsp

9. Can the authors provide details on the reasons why the EPAR was analysed both in print and on-screen?

- The documents are available to read both on screen and as a printed document. We wished to test the readability of both forms of the document.

10. In the discussion I would suggest to add a paragraph with the study limitations. I would also ask the authors to express their view in a more explicit way about whether they would recommend the approach described in the present paper to be applied by the EMA as a new standard for producing better EPARs. What about implications in terms of time and costs?

- We have added a 'Limitations' section in the Discussion as mentioned above in Point 24.
- In answer to the second part of this point, we are not recommending applying our approach in this paper as a new standard for producing better EPARs. We agree that the time and cost implications would be significant. The message from this study is designed to be that user testing can help to improve the general layout and content of EPAR summaries.

11. The authors reported under competing interests that they funded Luto Research which undertook the testing. Can they explain whether this is an intellectual or financial competing interest?

- We feel it is not for us as authors to interpret any competing interests – it is for the reader to decide, based on the information given.

VERSION 2 – REVIEW

REVIEWER	Katri Hämeen-Anttila, PhD Development Manager Finnish Medicines Agency FIMEA Finland
REVIEW RETURNED	25-Jun-2013

THE STUDY	I would prefer more detailed description of the data analysis (both quantitative and qualitative), however, I understand the word limit restrictive.
GENERAL COMMENTS	The authors have taken into account well the comments for the first version of the manuscript and it is much improved. I am happy to recommend accepting this manuscript for publication after some minor, mainly stylistic issues some of which are related to the BMJ Open instructions for authors: - instructions allow for up to five figures and tables – the manuscript includes now six figures, five tables and a Box (which is not recognized in the instructions). I suggest moving some of the tables and figures as supplementary data. - table 3 is not referred to in the text - all quotations should be in italics (page 12, lines 24–25) - I would prefer more detailed description of the data analysis (both quantitative and qualitative), however, I understand the word limit restrictive.

VERSION 2 – AUTHOR RESPONSE

Reviewer: Katri Hämeen-Anttila, PhD: Thank you for the opportunity to review this revised manuscript. The authors have taken into account well the comments for the first version of the manuscript and it is much improved.
- Thank you, we are pleased that our revised draft is much improved.

I am happy to recommend accepting this manuscript for publication after some minor, mainly stylistic issues some of which are related to the BMJ Open instructions for authors:

1.Instructions allow for up to five figures and tables – the manuscript includes now six figures, five tables and a Box (which is not recognized in the instructions). I suggest moving some of the tables and figures as supplementary data.
- We have not changed the number of figures and tables, in the light of the editor's note above.

2.Table 3 is not referred to in the text
- Reference is now made in the text on page 11

3. All quotations should be in italics (page 12, lines 24–25)
- This quotation is now in italics

4. I would prefer more detailed description of the data analysis (both quantitative and qualitative), however, I understand the word limit restrictive.
- We feel that the description of the data analysis (for both the quantitative and qualitative data) is sufficient. Readers are able to consult the papers we reference throughout the Introduction and Methods, to find more detailed descriptions.