PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Defining Acute Flares in Knee Osteoarthritis: A Systematic Review
AUTHORS	Parry, Emma; Thomas, M; Peat, George

VERSION 1 – REVIEW

REVIEWER	Jodie McClelland La Trobe University, Australia
REVIEW RETURNED	07-Nov-2017

GENERAL COMMENTS	Thank you for the opportunity to review this manuscript describing a systematic narrative review on the definition of acute flares in knee osteoarthritis. This research is interesting and provides the basis for an important discussion about how we talk about and classify exacerbations of knee osteoarthritis and I recommend it for publication. However I am seeking clarification on the some aspects of the manuscript that I have outlined below: Abstract: The term 'flare trials' needs further definition. I appreciate that you have described this clearly in the manuscript, but it is not clear what you mean by this term in the abstract. Limitations: Thank you for clearly stating the limitations of the search strategy. I have some concerns about why these decisions were made. Why were these broader search terms not included in the search strategy when the stated advantage of and need for this study is for a 'broad search of the literature'? Although these terms are not the most typical of those used to describe exacerbations of OA, it is likely that some studies were not identified by omitting them. This may be one of the reasons why a reasonably large number of studies were identified through reference searching. If the authors choose to proceed without broadening the search terms, then a stronger justification must be presented. Methods: It is only from the title of the manuscript that it is clear that you are interested in the definition of exacerbations of knee OA - please include clarification of the location of OA (knee) in the statement of inclusion criteria. Line 120 - Please clarify what you mean by 'with or without classification criteria based on measurement'. This appears to require a separate statement. Line 167 - This statement suggests that the duration of flare is somewhat equivalent to a withdrawal period, which is not the case. Please clarify this statement. Results Line 211 - Please clarify that by size the authors are referring to the number of participants or sample size. Discussion - This is a clear, concise and appropriate discussion that refers to literature discussing similar concepts in other populations.

REVIEWER	Caroline Terwee VU University Medical Center, Department of Epidemiology and

	Biostatistics, Amsterdam, the Netherlands
REVIEW RETURNED	13-Dec-2017

GENERAL COMMENTS	Comments to the authors This is a well-performed and clearly described systematic review of definitions of acute flare in knee osteoarthritis. It was my pleasure to review this paper. I only have a few minor questions and comments. • I would be interested to see which of the identified definitions were based on some kind of consensus procedure. Have you extracted that information and would it be possible to report this in the manuscript? • I did not understand the description on page 22 of the study of Marty et al. who validated a set of criteria, using logistic regression analysis….? Could you explain a bit more what was done in this study? • In Table 2, the difference between the third and fourth column is not really clear to me (both columns contain cut-offs, which is a bit confusing). • I wonder if it would be interesting (or of added value) to include the full definition of acute flare provided in the included studies? • In the discussion you state that how challenges for achieving consensus are resolved depends partly on whether the goal is a shared definition for reproducible and comparative research or for identifying these phenomena in routine practice. I do not see yet how that would be different?

REVIEWER	Cathy Chapple PhD School of Physiotherapy, University of Otago, New Zealand
REVIEW RETURNED	14-Dec-2017

GENERAL COMMENTS	This is a well conducted and clearly reported review. It complies with and reports against the PRISMA guidelines. The purpose was outlined and justified with reference to recent research and an identified gap in the literature. The search strategy and retrieval of articles is comprehensive. Presentation of results allowed for easy interpretation of findings. It would have been helpful to see the drug that was withdrawn in the withdrawal studies. For example the type of medication may have influenced the extent of flare-up, and wash-out period to flare-up. It may also have indicated a possible mechanism for the flare-up e.g. inflammation. The discussion was perhaps a little over cautious. There was concurrence of findings e.g. 8/10 studies used >40mm of change in pain as their threshold for defining a flare-up. This was worthy of highlighting in the discussion, perhaps with reference to minimally important clinical/detectable difference? Although not the purpose of the study some consideration of the possible mechanisms influencing flare-ups would have been interesting. Also a recognition of the heterogeneity of OA, the natural cycle of symptoms etc as factors which could potentially influence flare-up, and would therefore need to be considered in any definition of flare-up. It appears that the non-withdrawal trials were less specific about defining flare-ups. This could be highlighted with a recommendation to address this in future trials? P.43 Line 56 The authors have identified a recent similar review Cross et al 2017 (their ref 98). The reference is incomplete in the reference list. Given the similarity of the purposes of the two studies more discussion/comparison between the two studies would

	strengthen/add support for the current study, as well as highlighting the differences.

REVIEWER	Giovanni E. Ferreira School of Public Health, University of Sydney, Australia
REVIEW RETURNED	20-Dec-2017

GENERAL COMMENTS	This manuscript is a systematic review with a descriptive synthesis approach that aimed to describe the concept of symptom flare among a variety of study designs looking for a common, shared definition for research and clinical application. To answer this question, several databases were searched and both studies with "flare designs" as well as RCTs and observational studies were used. As pointed by the authors, most definitions of symptom flare are not consistent in the literature and the criteria used to determine whether or not a flare has occured remains uncertain. My main concern is with the overly simplistic way data was handled and described only in a descriptive manner. At the end of the results section, it is not clear what set of characteristics seem to be truly relevant. This is briefly highlighted in the first paragraph of the discussion, but not in enough detail, including suggestions on how future studies can adequately address this research question. Furthermore, among the 69 included studies, only two of them tried to derivate a prediction model to define what a symptom flare was. The other 67 just replicated common-sense and ad-hoc definitions. As the number of studies with an objective approach to answering the question "what characterises a symptom flare in knee osteoarthritis" was very limited, a reasonable alternative to improve data synthesis process is to perform a qualitative synthesis of results using a thematic synthesis approach (you can refer to Thomas & Hardem. BMC Med Res Methodol. 2008; 8: 45 for further guidance). Grouping terms into themes using coding and absctraction procceses can be a suitable alternative for this study provide a more objective answer to the proposed aims of the study. Please see some other minor comments: Line 134 Data Collection - This section can be merged with "literature sources and study selection" because it basically described the process of study selection. Also, authors have expatiated a bit on this topic. I would suggest a more direct writing style for this section. Line 163-164 - Data extraction - The second reviewer did not actively participate in the whole proccess of data extraction. This reduces the quality of a systematic review according to the AMSTAR checklist; therefore, it needs to be highlighted as a study limitation. Line 212 Add "sample" before "size"; add "participants" after "15-6085" Table 1. I would suggest authors to improve the standardisation of terms to describe the setting where the study was conducted. Terms like "community", "hospital", "primary care" were well applied, but the description of "clinical centres", "outpatients", "investigative sites" could be improved. Table 1 (page 19, line 33-39) - what is "au"? Please revise.

	Line 241 - this sentence has to be moved to the paragraph below Lines 244-252 - As these studies were the only trying to derivate/validate a prediction model on what a flare is, they should be describe in more detail. Hierarchically, they are more important than other studies that used arbitrary criteria. These studies should be acknowledged in the discussion as well. Page 41 - lines 9-14 - How many studies used each one of these single items? Page 41 - lines 35-37: "For studies using a drug withdrawal design the duration of the washout period differed between studies, ranging from 2-15 days" This seems more pertinent to the study design rather than to a relevant feature of a symptom flare-up.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

1.1 Abstract: The term 'flare trials' needs further definition

Response: This has now been expanded.
Revision: “The concept of OA flare appears in the medical literature but most often in the context of flare design trials (pain increases observed after stopping usual treatment).”

1.2 Limitations: Why were these broader search terms not included in the search strategy when the stated advantage of and need for this study is for a 'broad search of the literature'? Although these terms are not the most typical of those used to describe exacerbations of OA, it is likely that some studies were not identified by omitting them. This may be one of the reasons why a reasonably large number of studies were identified through reference searching. If the authors choose to proceed without broadening the search terms, then a stronger justification must be presented.

Response: Thank you for this comment. We appreciate that a thorough search strategy is crucial to the quality of the systematic review process. It is not uncommon for a high proportion of included articles to be identified through “snowball” methods such as screening reference lists (Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. BMJ;331:1064-5), and this was the case in our study, even after searching 11 databases. By comparison with the search strategy by Cross et al. 2017 we feel our search strategy was comprehensive yet efficient – returning 69 included articles compared with 23. While terms such as “attack” and “episode” may be used to describe flares in osteoarthritis, it was clear to us that the vast majority of ‘hits’ related to these terms would instead be picking up comorbidities (e.g. gout) and episodes of care.
Revision: Strengthened justification in discussion as follows:
“….although these terms appeared often to relate to comorbidities and other phenomena (e.g. episodes of care) and would therefore have led to a less efficient search strategy than relying on snowball references….. In comparison with Cross et al[21] our search strategy appeared comprehensive yet efficient – returning 69 included articles compared with 23.”

1.3 please include clarification of the location of OA (knee) in the statement of inclusion criteria.

Response: Inclusion criteria on line 121 state “flare-up of knee OA” but we recognise that some studies included a mixed OA population, typically knee or hip OA.
Revision: Added “Studies that included a mixed OA population (e.g. knee or hip OA) and did not separately report knee-specific findings were included.”

1.4 Line 120 - Please clarify what you mean by 'with or without classification criteria based on measurement'. This appears to require a separate statement.

Response: Clarified.
Revision: Added “Studies were included in the final full text peer review if they contained a description or definition of an acute exacerbation or flare-up of knee OA in human adults (18 years or over) in the general population, primary care or hospital settings. Studies were included even if their description was not based on clear measurement criteria (e.g. stating a ‘significant increase in pain’ but not what amount of change on a pain score this would equate to).”

1.5 Line 167 - This statement suggests that the duration of flare is somewhat equivalent to a withdrawal period, which is not the case. Please clarify this statement.

Response: Amended for clarity
Revision: Amended to “duration of flare (for flare design trials we extracted the duration of the withdrawal period for comparison)”

1.6 Results Line 211 - Please clarify that by size the authors are referring to the number of participants or sample size.

Response: Rephrased for clarity
Revision: “The number of participants in each study ranged from 15-6085[2, 3]”

Reviewer 2

2.1. which of the identified definitions were based on some kind of consensus procedure. Have you extracted that information and would it be possible to report this in the manuscript?

Response: The rationale behind each definition used was sought and described in Table 2 if presented in the corresponding study. None of the definitions appeared to be based on a consensus procedure.
Revision: Inserted in results “None of the definitions were based on a consensus procedure.”

2.2. I did not understand the description on page 22 of the study of Marty et al. who validated a set of criteria, using logistic regression analysis….? Could you explain a bit more what was done in this study?

Response: Apologies for this, we have provided further information.
Revision: Added “…a diagnostic tool for knee OA flares. Potential factors associated with flare-ups were identified, for example, knee swelling and the authors used a…”

2.3. In Table 2, the difference between the third and fourth column is not really clear to me (both columns contain cut-offs, which is a bit confusing).

Response: Apologies for this. Column 3 refers to the amount of change (increase) in pain that was required as part of the study’s definition of a flare. Column 4 refers to the minimum absolute level of pain experienced during a flare. For example, a patient whose pain increased from 0 to 3 out of 10 would experience a ‘significant increase’ in pain (greater than 2-point increase) but still only be experiencing relatively mild pain (i.e. below a minimum threshold of 4 out of 10 for example).
Revision: Column headings amended to “Minimum amount of change in symptoms/signs” and “Minimum absolute level of symptoms/signs”.

2.4. I wonder if it would be interesting (or of added value) to include the full definition of acute flare provided in the included studies?

Response: Many thanks for this suggestion. It is something we considered but discounted as it adds little value to synthesis.
Revision: None proposed.

2.5. In the discussion you state that how challenges for achieving consensus are resolved depends partly on whether the goal is a shared definition for reproducible and comparative research or for identifying these phenomena in routine practice. I do not see yet how that would be different?

Response: Rephrased for clarity
Revision: Each potential cardinal feature of OA flare presents different challenges for achieving consensus. The goal of an agreed composite definition is to facilitate both reproducible and comparable research, whilst enabling more consistent recognition and identification of these phenomena in routine practice.

Reviewer 3

3.1. It would have been helpful to see the drug that was withdrawn in the withdrawal studies. For example the type of medication may have influenced the extent of flare-up, and wash-out period to flare-up. It may also have indicated a possible mechanism for the flare-up e.g. inflammation.

Response: It is an interesting point. However, flare design trials withdraw potential participants from their usual treatment and this is not always clearly reported. We would respectfully suggest that this is a secondary question that would be better addressed in a different study.
Revision: No revision.

3.2. The discussion was perhaps a little over cautious. There was concurrence of findings e.g. 8/10 studies used >40mm of change in pain as their threshold for defining a flare-up. This was worthy of highlighting in the discussion, perhaps with reference to minimally important clinical/detectable difference?

Response: Thank you for this comment. The discussion has been amended to highlight this.
Revision: Inserted “There was general concurrence with the minimum threshold that was adopted, for example, 40mm on a 0-100mm scale and this may indicate the potential level of minimally important clinical difference”.

3.3. Although not the purpose of the study some consideration of the possible mechanisms influencing flare-ups would have been interesting

Response: In undertaking research in this field we recognise the importance of better understanding mechanisms of flares but have also been criticised for speculating on this without adequate evidence from the study under question.
Revision: No revision

3.4. Also a recognition of the heterogeneity of OA, the natural cycle of symptoms etc as factors which could potentially influence flare-up, and would therefore need to be considered in any definition of flare-up.

Response: Thank you for this comment. We also feel the heterogeneity of OA may potentially influence flare-up behaviour and have added a statement in the discussion. In paragraph two of the discussion we have already considered the importance of day-to-day variability.
Revision: Add 'The heterogeneity of OA should also be considered in any definition of a flare-up.' To the first paragraph of the discussion.

3.5. It appears that the non-withdrawal trials were less specific about defining flare-ups. This could be highlighted with a recommendation to address this in future trials?

Response: We feel the non-withdrawal trials are more likely to reflect 'naturally' occurring flares and the uncertainty surrounding their diagnosis in clinical practice, which may explain why these studies are less specific with regards to their definitions. The withdrawal trials require more specific definitions to ensure methodological rigour within their studies. Flare or exacerbation definitions which are widely used in other chronic diseases, for example, COPD are not always specific and it may be the case that a less specific definition will be more helpful for defining knee OA flares.
Revision: No revision as the discussion highlights the need for further consensus exercises to reach a common shared definition of knee OA flares.

3.6. P.43 Line 56 The authors have identified a recent similar review Cross et al 2017 (their ref 98). The reference is incomplete in the reference list. Given the similarity of the purposes of the two studies more discussion/comparison between the two studies would strengthen/add support for the current study, as well as highlighting the differences.

Response: Please see response to editor's comment
Revision: Reference has been completed. Revisions as per E2 above.

Reviewer 4
4.1. At the end of the results section, it is not clear what set of characteristics seem to be truly relevant. This is briefly highlighted in the first paragraph of the discussion, but not in enough detail, including suggestions on how future studies can adequately address this research question.

Response: Thank you for this comment. We feel that the next step to achieving a widely accepted and usable definition is a consensus exercise and we have highlighted that the cardinal features identified in this review could be used as a starting point.
Revision: Added to end of discussion "The cardinal features described in this review; onset/worsening of symptoms and signs, attainment of a minimum symptom threshold during flare, speed of onset/worsening, and duration of elevated symptoms/signs will hopefully help start this discussion."

4.2. improve data synthesis process is to perform a qualitative synthesis of results using a thematic synthesis approach

Response: We have given detailed consideration to this suggestion. The thematic synthesis advocated by Thomas & Hardem. BMC Med Res Methodol. 2008; 8: 45 offers an approach to the detailed analysis of primary qualitative research. We respectfully wish not to adopt this alternative approach for the following reasons. We appreciate that the flare-definitions are qualitative, but our objective was not to perform a thematic analysis of primary qualitative research and to do so now would be to undertake a different systematic review with different objectives. Also, the predominant use of brief, common-sense and ad-hoc definitions make a formalised thematic analysis difficult to perform in any meaningful way that would enhance the presentation of the results.
Revision: No revision

4.3. Line 134 Data Collection - This section can be merged with "literature sources and study selection" because it basically described the process of study selection. Also, authors have expatiated a bit on this topic. I would suggest a more direct writing style for this section.

Response: Amended as requested
Revision: Removed separate heading for data collection and edited down this section and data extraction

4.4. Line 163-164 - Data extraction - The second reviewer did not actively participate in the whole process of data extraction. This reduces the quality of a systematic review according to the AMSTAR checklist; therefore, it needs to be highlighted as a study limitation.

Response: Please response to E3 above
Revision: As per E3 above.

4.5. Line 212 Add "sample" before "size"; add "participants" after "15-6085"

Response: Please see response to 1.6 above
Revision: As per 1.6 above

4.6. Table 1. I would suggest authors to improve the standardisation of terms to describe the setting where the study was conducted. Terms like "community", "hospital", "primary care" were well applied, but the description of "clinical centres", "outpatients", "investigative sites" could be improved.

Response: The terms used to describe the study setting were taken verbatim from the original studies. We have been cautious not to over interpret these terms and have decided to leave the terms as presented in the corresponding texts.
Revision: No revision

4.7. Table 1 (page 19, line 33-39) - what is "au"? Please revise.

Response: It was an abbreviation for Australia
Revision: Changed to ‘Australia’

4.8. Line 241 - this sentence has to be moved to the paragraph below

Response: Thank you for highlighting this.
Revision: Sentence moved to paragraph below.

4.9. Lines 244-252 - As these studies were the only trying to derivate/validate a prediction model on what a flare is, they should be describe in more detail. Hierarchically, they are more important than other studies that used arbitrary criteria. These studies should be acknowledged in the discussion as well.

Response: Thank you for this comment. We agree that the studies by Marty et al and Scott-Lennox deserve more attention.
Revision: Moved paragraphs on Marty et al and Scott-Lennox et al to beginning of section discussing rationale for flare definitions. Included more detail on these two studies “The study by Marty et al[20] was the only study specifically designed to validate a diagnostic tool for knee OA flares. Potential factors associated with flare-ups were identified, for example, knee swelling and the authors used a logistic regression analysis to assign a weight to each of the items identified. A flare up score was

determined using a general practitioner database and this was then validated using a rheumatologist database. Pain was not included in the final model".

"Scott-Lennox et al[58] sought to test whether four measures for flare intensity (patient's self-assessment of pain scores, physician's assessment of pain scores, patient's global OA assessment and physician's global OA assessment) could be combined to form a reliable and valid index using data from an RCT using a confirmatory factor analysis. The authors produced three flare intensity groups (low, moderate and severe) and highlighted how these could be used to examine treatment effects".

Inserted into paragraph two of the discussion "However, the study by Marty et al[20] and Scott-Lennox et al[58] were the only two studies we found that had attempted to derive and/or validate a prediction model for OA flares. Interestingly their approaches have not been widely adopted which suggests the complexity of reaching a widely accepted model".

4.10. Page 41 - lines 9-14 - How many studies used each one of these single items?

Response: We have now reported the number of studies using only single item, only multi-item and both single and multi-item tools.
Revision: Added the following to pg 33 "Thirty-four studies used only single item measurement tools[29-32, 34, 36-45, 47, 49, 50, 52, 54, 57, 58, 60, 61, 63-65, 75-79, 92, 93], 5 used multi-item[33, 48, 53, 55, 62] and 5 used both single and multi-item tools[26, 28, 35, 91, 95]".

4.11. Page 41 - lines 35-37: "For studies using a drug withdrawal design the duration of the washout period differed between studies, ranging from 2-15 days" This seems more pertinent to the study design rather than to a relevant feature of a symptom flare-up.

Response: Thank you for this comment. On reflection, it is confusing to include this in the results section as it describes part of the study design. It may be relevant as a trigger for symptom flare but that is beyond the scope of our study.
Revision: Removed from the results section; "For studies using a drug withdrawal design the duration of the washout period differed between studies, ranging from 2-15 days"

VERSION 2 – REVIEW

REVIEWER	Cathy Chapple PhD School Of Physiotherapy, University of Otago, New Zealand
REVIEW RETURNED	01-Mar-2018

GENERAL COMMENTS	Thank-you for addressing reviewers comments, either by adapting the manuscript or by justifying "no revision" I feel the changes have improved the clarity of reporting your systematic review, and made it more straightforward for a non-specialist to read and absorb the message. The response to the comment about inclusion of the Cross systematic review is acknowledged, but lines 86-87 seem a slightly clumsy way of acknowledging it. Perhaps it is less about justifying the novel contribution of your paper, and more that this is clearly an interesting and important area for consideration (in intro). And in the discussion that there is added strength in the findings as two reviews are addressing this area, and your findings will extend/expand those of the recent review. Line 170 could do with a full stop after the second use of word

	"studies". New sentence begins with "Frequencies" to highlight you are talking about the features of all the studies. The fact that function was not identified as a cardinal feature is interesting, but only relatively few studies evaluated function, and those that did looked at walking. Except for those studies that included composite scores, factors such as walking up/down stairs, getting in/out of chair would not have been included. These are often the functions where pain would reveal itself first. Not sure you can add anything at this stage but it is interesting to note.

REVIEWER	Giovanni Ferreira Musculoskeletal Health Sydney, School of Public Health, University of Sydney, Sydney, Australia
REVIEW RETURNED	05-Mar-2018

GENERAL COMMENTS	I thank the authors for having acknowledged the several suggestions made by the review team. I commend their effort for having amended the manuscript in such a way that its quality has definitely improved. I am willing to recommend this manuscript for publication contingent upon minor corrections. METHODS 1. The data extraction section is not written in a clear way with regards to the disclosure of the conventional steps of a systematic review that weren't followed. Instead of writing "Extraction for every tenth article was independently checked (MJT)." (line 156), please make it clearer that the second reviewer has not been involved in the process of data extraction. DISCUSSION 2. The manuscript made it clear that more research is needed in order to understand what clinical features truly constitute an acute flare, hence I believe it would be important for this manuscript to highlight how future research should address this issue. Specifically, describe what research methods/study designs/analysis would be relevant to provide relevant insights and to contribute this field of knowledge.

VERSION 2 – AUTHOR RESPONSE

Reviewer 3

3.1 The response to the comment about inclusion of the Cross systematic review is acknowledged, but lines 86-87 seem a slightly clumsy way of acknowledging it. Perhaps it is less about justifying the novel contribution of your paper, and more that this is clearly an interesting and important area for consideration (in intro). And in the discussion that there is added strength in the findings as two reviews are addressing this area, and your findings will extend/expand those of the recent review.

Response: Thank you for highlighting this. We have edited the manuscript to acknowledge the recent review and the contribution our findings make.

Revision: Deleted ‘A review addressing similar aims but not registered on PROSPERO came to our attention when it was published while we were drafting our manuscript [21]. In principle and upon comparing the respective findings of both reviews, we felt our review could justify making an original contribution.’ From introduction and ‘The recently published review that sought to define flare-ups in in

hip and knee OA only yielded 23 studies and four of the included studies did not contain clear definitions for a flare-up [21]. Furthermore,' from discussion.

Added to discussion 'We feel that our review expands on the findings of the Cross et al review and adds strength to this important area'.

3.3 Line 170 could do with a full stop after the second use of word "studies". New sentence begins with "Frequencies" to highlight you are talking about the features of all the studies.

Response: Thank you for highlighting this.

Revision: Amended as suggested.

3.4 The fact that function was not identified as a cardinal feature is interesting, but only relatively few studies evaluated function, and those that did looked at walking. Except for those studies that included composite scores, factors such as walking up/down stairs, getting in/out of chair would not have been included. These are often the functions where pain would reveal itself first. Not sure you can add anything at this stage but it is interesting to note.

Response: Many thanks for the above comment which we agree with.

Revision: None made.

Reviewer 4

4.1 Methods: The data extraction section is not written in a clear way with regards to the disclosure of the conventional steps of a systematic review that weren't followed. Instead of writing "Extraction for every tenth article was independently checked (MJT)." (line 156), please make it clearer that the second reviewer has not been involved in the process of data extraction.

Response: Thank you for highlighting this. We have opted to delete the sentence to avoid confusion.

Revision: Deleted 'Extraction for every tenth article was independently checked (MJT).'

4.2 Discussion: The manuscript made it clear that more research is needed in order to understand what clinical features truly constitute an acute flare, hence I believe it would be important for this manuscript to highlight how future research should address this issue. Specifically, describe what research methods/study designs/analysis would be relevant to provide relevant insights and to contribute this field of knowledge.

Response: The last paragraph of the discussion highlights the need for a consensus exercise in order to reach an agreed definition amongst key stakeholders. Further research to gain insights into 'naturally' occurring flares rather than those that are drug induced could be achieved through observational studies using repeated measures.

Revision: Inserted the following sentence into last paragraph of discussion 'Furthermore, observational studies with repeated measures could give an important insight into the nature of these phenomena.'

VERSION 3 – REVIEW

REVIEWER	Giovanni Ferreira University of Sydney, Australia
REVIEW RETURNED	14-Apr-2018

GENERAL COMMENTS	Thanks again for the opportunity of reviewing this manuscript. All questions were adequately addressed by the authors. Having no further comments to add, I am happy to recommend this manuscript to be accepted for publication.