PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Can the appropriateness of eyecare be measured through cross-sectional retrospective patient record review in eyecare practices in Australia? The iCareTrack feasibility study
AUTHORS	Ho, Kam; Rahardjo, Dian; Stapleton, Fiona; Wiles, Louise; Hibbert, Peter; White, Andrew; Hayen, Andrew; Jalbert, Isabelle

VERSION 1 – REVIEW

REVIEWER	Bárbara Antunes Centro de Estudos e Investigação em Saúde da Universidade de Coimbra, Portugal
REVIEW RETURNED	19-Aug-2018

GENERAL COMMENTS	Can the appropriateness of eyecare be measured through retrospective patient record review in eyecare practices? The iCareTrack feasibility study This is a well written report. I have a few minor comments. I hope the authors find them useful. 1. In the abstract it reads: “Primary Outcome Measure: … The outcome for the main trial will be the percentage of encounters at which the appropriate eyecare is received.” 2. Also, it would be helpful for the reader to have a definition of “encounters at which the appropriate eyecare is received” is this the number of consults each patient has at the site of care? 3. Related with the above point, was it not possible to look into a sample of patient records and determine the % of encounters during which appropriate eyecare was received and include those data in the onsite retrospective record review of eyecare practices designed for the feasibility study? Could this not have been an aim of the feasibility study? Would it not help to inform the main study? 4. In the article summary section the last bullet point reads: “The study findings support the conduct of a larger trial to determine the percentage of eyecare encounters at which appropriate care is received”. In what way? Could authors specify? Typo Box 1: Eyecare provider is an individual who provides ("s" is missing) There is no trial registration information reported. Was this study registered?

	There is no mention of any reporting checklist.

REVIEWER	Patricia Hrynchak University of Waterloo, School of Optometry and Vision Science, Canada
REVIEW RETURNED	24-Oct-2018

GENERAL COMMENTS	Can the appropriateness of eyecare be measured through retrospective patient record review in eyecare practices? The iCareTrack feasibility study This is a very interesting study from a public health policy perspective and a quality assurance perspective. Quality assurance falls under the domain of the colleges of optometrists in Canada and a method such as this one would be useful in doing retrospective record reviews to determine quality practice at the individual provider level. Minor points: 1. There should be commas between the reference numbers so that 1,2 is clearly different from 12. 2. In the definitions the terms client and patient are used. One or the other should be used for consistency and clarity. 3. There are small grammatical errors throughout which should be easily fixed. 4. The definition of evidence-based practice normally includes patient values. That aspect seems to be missing in the definition. 5. Page 7 second to last sentence in first paragraph – perhaps you would want to optimize spending rather than maximize it? 6. Reference 8, consider the appropriateness of referencing a paper that is not available. 7. Page 9' what is the difference between having risk factors for glaucoma and being a glaucoma suspect? 8. Page 9, add "type 2" before diabetes mellitus 9. Page 9, sentence starting with "Pregnant patients" needs to be made into 2 sentences. 10. Were there any practices that refused to participate in the study? This should be stated e.g., were there 20 practices approached and 8 agreed? Major points: 1. The purpose of the study is clearly to identify the feasibility of retrospective onsite file review not to determine the appropriateness of care. Given the significant problems with validity of the data (e.g., different indicators used throughout the study) those results should not be in the abstract rather more information about the feasibility issues. 2. The term "random sample" is used repeatedly in the paper. An explanation of how randomization was done is needed.

	3. Were the guidelines that were selected appraised for quality? How was that done? Not all guidelines are of good quality i.e., COS guidelines for periodic eye examinations. 4. This is not relevant for this paper but how are you going to deal with the fact that some of the guidelines have been updated since this study was done? 5. At the University of Waterloo, in order to use patient records for retrospective research purposes patients need to consent in advance. Is that not the case in your context? 6. It would be good to calculate a kappa statistic for inter-rater reliability of data abstraction 7. On page 14 you say the sampling continued until the required number of records was found but that did not occur as the required number was 10 and the average for DM was 4. Can you explain this? Also the preventative care was 23 median but you aimed at 10; why was that the case? 8. For the "randomly sampled individual record" on page 14 do you mean the alphabet method? I think the types of sampling need to be more clear so the reader is not confused as to what actually happened. 9. Why was the appropriateness of preventative care only assessed in three practices but with more files than intended in the initial protocol? An explanation of why this happened and the glaucoma and diabetic eyecare in only 7 of the 8. 10. I think the results of appropriateness should not be in a figure. The results are not valid due to the study limitations and therefore should not be emphasized. 11. Discussion: if feasibility is not condition specific, why were only 3 practices used for preventative care? 12. What decisions were made from this study that informed the larger study? Please address this topic.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Bárbara Antunes
Institution and Country: Centro de Estudos e Investigação em Saúde da Universidade de Coimbra, Portugal
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below
Can the appropriateness of eyecare be measured through retrospective patient record review in eyecare practices? The iCareTrack feasibility study

This is a well written report. I have a few minor comments. I hope the authors find them useful.

1. In the abstract it reads: "Primary Outcome Measure: … The outcome for the main trial will be the percentage of encounters at which the appropriate eyecare is received."
Response: The sentence has been modified to address how the feasibility study informs the main study. 'A secondary outcome was the percentage of practitioner-patient encounters at which appropriate eyecare was received.' (Page 3)

2. Also, it would be helpful for the reader to have a definition of “encounters at which the appropriate eyecare is received” is this the number of consults each patient has at the site of care?
Response: This is the percentage of consults at which patients receive appropriate eyecare. The sentence in the abstract has been clarified to state ‘the percentage of practitioner-patient encounters at which appropriate eyecare was received.’ (Page 3)

3. Related with the above point, was it not possible to look into a sample of patient records and determine the % of encounters during which appropriate eyecare was received and include those data in the onsite retrospective record review of eyecare practices designed for the feasibility study? Could this not have been an aim of the feasibility study? Would it not help to inform the main study?

Response: We agree with the reviewer that the percentage of encounters during which appropriate eyecare was received is also a secondary outcome for the feasibility study. We have reworded the sentence as indicated above. ‘A secondary outcome was the percentage of practitioner-patient encounters at which appropriate eyecare was received.’ (Page 3)

4. In the article summary section the last bullet point reads: “The study findings support the conduct of a larger trial to determine the percentage of eyecare encounters at which appropriate care is received”. In what way? Could authors specify?

Response: We added the following sentence. ‘This study identified the best patient record sampling methods, potential issues associated with recruiting eyecare practices, and with accessing, extracting, recording and analysing clinical records.’ (Page 4)

Typo Box 1: Eyecare provider is an individual who provides ("s" is missing)
Response: We thank the reviewer for pointing out the typo and we have changed ‘provide’ to ‘provides’. (Page 6)

There is no trial registration information reported. Was this study registered?

Response: This study was a cross-sectional retrospective patient record review, so this study was not registered.

There is no mention of any reporting checklist.

Response: A completed STROBE checklist with page number indicating the location of each item is attached.

Reviewer: 2

This is a very interesting study from a public health policy perspective and a quality assurance perspective. Quality assurance falls under the domain of the colleges of optometrists in Canada and a method such as this one would be useful in doing retrospective record reviews to determine quality practice at the individual provider level.

Minor points:

1. There should be commas between the reference numbers so that 1,2 is clearly different from 12.
   Response: We agreed reviewer's comment and commas were added between references numbers between 1 and 2 (page 5), between 4 and 5 (page 5), between 15 and 20-22 (page 7), between 18 and 30 (page 9), and between 39 and 40 (Page 16).

2. In the definitions the terms client and patient are used. One or the other should be used for consistency and clarity.
   Response: 'Client' was replaced by 'patient' for consistency in the main text (page 5) and Box 1 (page 6).

3. There are small grammatical errors throughout which should be easily fixed.
   Response: The revised manuscript has been carefully checked for errors

4. The definition of evidence-based practice normally includes patient values. That aspect seems to be missing in the definition.
   Response: We thank reviewer for pointing out this omission in the definition. 'and patient's preferences' are added to the definition of 'evidence-based care' in the Box 1 (Page 6).

5. Page 7 second to last sentence in first paragraph – perhaps you would want to optimize spending rather than maximize it?
   Response: We thank reviewer for the suggestion of using 'optimise' and the change has been made accordingly (Page 6).

6. Reference 8, consider the appropriateness of referencing a paper that is not available.
   Response: The unpublished reference was replaced by a published paper (Fung SS et al, BJO 2013) to support 'However, there is mounting evidence that such evidence-based clinical practice guidelines are not always adhered to or fully implemented in the clinical setting'. (Page 5)

7. Page 9' what is the difference between having risk factors for glaucoma and being a glaucoma suspect?
   Response: The following definitions were added to the manuscript (page 9). 'According to the Australian NHMRC guidelines, a glaucoma suspect is a person suspected of having glaucoma who has some but not all of the criteria required for a glaucoma diagnosis. They may have one or more of the following: suspicious optic disc, optic disc margin haemorrhage, occludable drainage angle, peripheral anterior synechiae or elevated intraocular pressure. A person at risk of glaucoma may be someone with a positive family history of glaucoma, or a history of chronic steroid use, or some other known risk factors for the disease.'

8. Page 9, add "type 2" before diabetes mellitus
   Response: 'type 2' is added before 'diabetes mellitus'. (Page 9)

9. Page 9, sentence starting with "Pregnant patients" needs to be made into 2 sentences.

Response: The sentence has been broken down into 2 sentences on page 9.
'Pregnant patients were excluded from all appropriateness assessments and patients with type 1 diabetes mellitus were excluded from the assessment of appropriateness of diabetic eyecare.'
And
'This was because different sets of clinical indicators were expected and the prevalence is likely to be too low to measure the appropriateness of eyecare for pregnant patients and patients with type 1 diabetes mellitus.'

10. Were there any practices that refused to participate in the study? This should be stated e.g., were there 20 practices approached and 8 agreed?
    Response: Nine practices were approached and eight agreed with one practice owner did not feel comfortable with giving access to the patient records. Modification was made as below:
    'Nine eyecare practices were invited and eight agreed to participate in the feasibility study. One practice refused to participate as the practice owner did not feel comfortable giving access to the patient records.'
    (Page 13)

Major points:

1. The purpose of the study is clearly to identify the feasibility of retrospective onsite file review not to determine the appropriateness of care. Given the significant problems with validity of the data (e.g., different indicators used throughout the study) those results should not be in the abstract rather more information about the feasibility issues.
    Response: The reviewer is correct in stating that the primary purpose of the study was to assess the feasibility of retrospective onsite review. However, we respectively disagree with this reviewer's suggestion to remove the appropriateness of care results. This suggestion is in contrast with the other reviewer's suggestion that the appropriateness of care results are an interesting secondary outcome that deserves inclusion in the abstract. Whilst Delphi 1st round indicators were used in this feasibility study, this does not invalidate the data: the indicators were not substantially changed between the first and final round and therefore, the impact of using early round indicators on the appropriateness of eyecare findings would be expected to be minimal.

2. The term "random sample" is used repeatedly in the paper. An explanation of how randomization was done is needed.
    Response: Patient records were randomly selected from the practice. The methods varied between different practices. A description of the random sampling method has been added at the end of the paragraph on page 12 as follows
    'Random sampling was achieved by using a random number generator. First, a range of number started from '1' was assigned to each patient or visit between 1st January 2013 and 31st December 2014. A list of random numbers was then generated and the records with the corresponding number was sampled.'

3. Were the guidelines that were selected appraised for quality? How was that done? Not all guidelines are of good quality i.e., COS guidelines for periodic eye examinations.

Response: The guidelines that were selected were not appraised for quality. However, the clinical indicators used in the main study were reviewed by the expert panels using the Delphi method to ensure the quality of the clinical indicators.

4. This is not relevant for this paper but how are you going to deal with the fact that some of the guidelines have been updated since this study was done?

Response: The latest version of clinical practice guidelines published prior to the study period (i.e. 1st January 2013) were used to allow time for the dissemination of the guidelines.

5. At the University of Waterloo, in order to use patient records for retrospective research purposes patients need to consent in advance. Is that not the case in your context?

Response: A waiver of patient consent was granted along with the ethics approval as described on page 13. This complies with National Privacy Principle 10.3 and NHMRC Chapter 2.3.6 "Qualifying or waving conditions for consent"[1]: This study involves minimal risk (that is already being faced) to both health care providers and participants, but the study provides information relevant to public health. This study can provide significant information and knowledge on whether the public receiving eye care compliant with evidence-based practice in Australia and cannot be achieved without access to records, which satisfies the requirements in Section A.1 of the NHMRC "Guidelines approved under Section 95A of the Privacy Act 1988"[2].

Reference:

1. National Health & Medical Research Council (NHMRC). National Statement on Ethical Conduct in Human Research 2007 (revised March 2014) Canberra, 2014. https://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/e72_national_statement_march_2014_140331.pdf

2. National Health & Medical Research Council (NHMRC). Guidelines approved under Section 95A of the Privacy Act 1988 Canberra, 2014. http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/pr2_guidelines_under_s95a_of_the_privacy_act_140311.pdf

6. It would be good to calculate a kappa statistic for inter-rater reliability of data abstraction

Response: Thank you. Kappa score yielded a substantial agreement between the two surveyors. We added 'Kappa score will be calculated to test the level of agreement between the two surveyors.' (Page 13) in the method and the result 'Substantial inter-rater agreement between the two surveyors was shown with a kappa score equal to 0.76 (95%CI 0.74, 0.78).' (Page 16)

7. On page 14 you say the sampling continued until the required number of records was found but that did not occur as the required number was 10 and the average for DM was 4. Can you explain this? Also the preventative care was 23 median but you aimed at 10; why was that the case?

Response: This is already explained on page 11 (methods) as follows 'It was initially intended that 10 records per condition per practice be reviewed, however in practice the number of records sampled varied based on the time available and the complexity of the on-site record review.' The following statement was added to clarify the methods 'Records of preventative care patients were first reviewed, followed by records of glaucoma and diabetic eyecare patients.' (page 11)
In addition, the following paragraph was added to the results section 'The purpose of this study was to explore potential issues associated with accessing, extracting, recording and analysing clinical records. Eligible records for preventative eyecare were easily identified, so time was preferentially allocated to reviewing of these over the other two

conditions, to identify the optimal sampling method. As a result, a median of 23 records (range: 21 to 50 records) were reviewed for preventative eyecare within the allowed time in three practices, which was more than the intended 10 records.' (page 14)

8. For the "randomly sampled individual record" on page 14 do you mean the alphabet method? I think the types of sampling need to be more clear so the reader is not confused as to what actually happened.

Response: The records were sampled based on the random number generated. This has been added at the end of page 11 to clarify the method.
'Random sampling was achieved by using a random number generator. First, a range of number started from '1' was assigned to each patient or visit between 1st January 2013 and 31st December 2014. A list of random numbers was then generated and the records with the corresponding number was sampled.' (Page 11)

9. Why was the appropriateness of preventative care only assessed in three practices but with more files than intended in the initial protocol? An explanation of why this happened and the glaucoma and diabetic eyecare in only 7 of the 8.

Response: Measurement of the appropriateness of preventative eyecare was demonstrated to be clearly feasible with a small number of practices, however, a larger number of practices were required to fully test the feasibility of measuring the appropriateness of care for the other two conditions. As a result, the time at subsequent practices was allocated to review of records for the other two conditions This is explained as shown above in point 7 and below.
'The purpose of this study was to explore potential issues associated with accessing, extracting, recording and analysing clinical records. Eligible records for preventative eyecare were easily identified, so time was preferentially allocated to reviewing of these over the other two conditions, to identify the optimal sampling method. As a result, a median of 23 records (range: 21 to 50 records) were reviewed for preventative eyecare within the allowed time in three practices, which was more than the intended 10 records.' (Page 14)

10. I think the results of appropriateness should not be in a figure. The results are not valid due to the study limitations and therefore should not be emphasized.

Response: These concerns have been addressed under point #1 above and are in contrast with comments from the other reviewer wishing for us to include measurement of appropriateness of eyecare as a secondary aim. We would like to retain the figure, as it provides interesting and useful information, and we feel that the limitations are clearly identified in the manuscript.

11. Discussion: if feasibility is not condition specific, why were only 3 practices used for preventative care?

Response: Measurement of the appropriateness of preventative eyecare was demonstrated to be clearly feasible with a small number of practices, however, a larger number of practices were required to fully test the feasibility of measuring the appropriateness of care for the other two conditions. As a result, the time at subsequent practices was allocated to review of records for the other two conditions.

12. What decisions were made from this study that informed the larger study? Please address this topic.

    Response: The following statement was already present in the discussion 'The most feasible and efficient sampling method that could be used in a majority of practices was to sample randomly by patient' (Page 17).
    The following sentence was added to address how this study informed the larger study.
    'Other recommendations derived from this feasibility study include the use of surveyors with an eyecare background, and a time allocation of 4 to 5 hours per practice for measurement of appropriateness of eyecare in three conditions.' (Page 17).

VERSION 2 – REVIEW

REVIEWER	Bárbara Antunes Centro de Estudos e Investigação em Saúde da Universidade de Coimbra
REVIEW RETURNED	07-Jan-2019

GENERAL COMMENTS	The paper is improved and all reviewers comments have been properly addressed.

REVIEWER	Patricia Hrynchak University of Waterloo, School of Optometry and Vision Science
REVIEW RETURNED	13-Dec-2018

GENERAL COMMENTS	All of my concerns and questions have been adequately addressed. Congratulations on producing an interesting paper.