PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Quality indicators for in-hospital geriatric co-management programmes: a systematic literature review and international Delphi study
AUTHORS	Van Grootven, Bastiaan; McNicoll, Lynn; Mendelson, Daniel; Friedman, Susan; Fagard, Katleen; Milisen, Koen; Flamaing, Johan; Deschodt, Mieke

VERSION 1 – REVIEW

REVIEWER	Simon Conroy University Hospitals of Leicester, UK
REVIEW RETURNED	20-Nov-2017

GENERAL COMMENTS	Nice paper with useful findings that can inform quality improvement and benchmarking. Well written (although the US spelling may not be in keeping with the house style, and numbers 0-9 should be written in full). Limitation acknowledged appropriately. Smallish sample size for the Delphi process and it is a shame that nurses and therapists were underrepresented – why? What was done to engage them?

REVIEWER	Victoria Tang University of California, San Francisco, United States
REVIEW RETURNED	02-Dec-2017

GENERAL COMMENTS	This manuscript is timely and highly relevant. As an increasing number of hospital systems are implementing geriatric co-management programs and the likes, it will be especially important to provide programs a starting point as to how to design and assess their programs. Additionally, this number will likely exponentially grow as the launching of the American College of Surgeon's Coalition for Quality in Geriatric Surgery (CQGS) Project will occur in 2019. The methodology strengths include the use of UCLA/RAND appropriateness method, which has been widely accepted as a means to assessing consensus in a Delphi study. Additionally, the development of an international group of geriatric co-management experts allows for a stronger validity to their assessment of the indicators developed from the systematic review. I agree the weakness lies within the lack of empirical evidence supporting the indicators that they have chosen, but as the authors have appropriately noted, there is a dearth of quality evidence and,

	because of this, expert opinion will need to be the first step to helping us develop much needed program assessment indicators. Table 1: In regards to the Age based characteristics, I'm assuming the categories are not mutually exclusive. Is that correct? The category of "Age <65 years" confused me on first pass. Some definitions would be helpful, (i.e., team meetings" and "transitional care")

REVIEWER	Alexander Joeris AO Foundation, Clinical Investigation and Documentation, Switzerland
REVIEW RETURNED	11-Dec-2017

GENERAL COMMENTS	I want to congratulate the authors for this tremendous work and effort they underwent for this project. I would like to encourage the authors to perform some minor revisions of the manuscript before it could be accepted for publication in BMJ Open: 1) Systematic literature review: the systematic literature review should be described in more details; additionally to the statement which databases were searched, the authors should either explain in the main manuscript or attach as a supplement in more details which search terms and combinations of search terms were used. Additionally, it would be of importance for the reader to explain in more detail how the 12'794 initially identified records were condensed down to 44 text articles for data extraction. 2) Selection of participants: The authors are asked to explain in more detail, why only experts from Europe and North America were selected. Although this fact is mentioned in the methodological considerations of the project and it is recommended to test the indicators in other regions, no rational is given why this was decided initially. Additionally I want to encourage the authors to give a reasoning why mot more experts were invited for the Delphi process. Although the authors highlight the fact that the majority of experts were medical doctors and only a few nurses and one manager, even within the medical doctors one orthopedic surgeon was included only. Looking at the reference list, there would have been a considerably number of orthopedic surgeons which could have participated potentially. As the acceptance of a consensus largely depends on the composition of the Delphi panel, the authors should give a reasoning for the decision to assemble the panel as it was. This especially, as in 87% of the articles the patient population of interest was "surgical". 3) Delphi study: it is not clear for the reader how the indicators in round 1 were compiled/re-worded/removed or added. It would be reasonable that the authors add a table as supplement which contain the indicators for round 1, which indicators were removed and which added for round 2. 4) Result section: Outcome indicators (p19 line 6): in fig 2 it is written that 17 outcome indicators are considered appropriate and feasible, in the text the authors state 16. Is this a pure reporting mistake or does the figure has to be interpreted differently? If so, please explain better, as otherwise the numbers are mis-leading.

	5) Discussion: p21, line 54ff: it is not surprisingly, but concerning that post-discharge follow-up outcomes were generally not considered feasible by the experts. This indicates the urgent need and importance to develop and implement post-discharge outcome measures which are thought to be feasible. 6) Methodological considerations: The authors state that the quality of the primary studies was generally poor. Is this a subjective statement or were measures taken to support this statement? If so, please describe/provide more information.

VERSION 1 – AUTHOR RESPONSE

Manuscript ID: bmjopen-2017-020617

Title: Quality indicators for in-hospital geriatric co-management programmes: a systematic literature review and international Delphi study

Comments on decision letter
Dear Editor in chief, Associate Editor, and Reviewers,
We are thankful for the constructive comments that we have received from the editor and reviewers. Please find below our respective revisions and explanatory comments. We are convinced that this has strengthened our manuscript and hope it will prove satisfactory. All the changes that we have made in the text have been highlighted using Track Changes.

Editorial requests
1) Please add a statement in the methods section clarifying whether written informed consent was obtained from participants. Please also briefly clarify in the methods why the study did not require approval from a local ethics committee.
- Response: We have added the following text in the methods section: “All participants consented to participate in the study via e-mail. Approval by a local ethics committee was not required as a Delphi study with healthcare professionals is not considered an experiment (Belgian law dated 7th May 2004 related to experiments on human people).”

Reviewer 1 (Simon Conroy)
1) Well written (although the US spelling may not be in keeping with the house style).
- Response: We have changed the text in adherence to the UK spelling.

2) Numbers 0-9 should be written in full.
- Response: We have changed the numbers 0 to 9 in the text as requested.

3) Smallish sample size for the Delphi process and it is a shame that nurses and therapists were underrepresented – why? What was done to engage them?
- Response: We acknowledge that the selection of participants is a key issue in Delphi studies as they are the instruments used to develop the quality indicators. However, to the best of our knowledge there are no guidelines to determine a minimum sample size for a Delphi study (Keeney et al. J Adv Nurs 2006;53(2):205-212). While we acknowledge that more participants lead to more data and more information, this does not mean that the results would be different. Also, we do not believe that the small number of non-medical professionals invalidates our results. The number of experts that we aimed to recruit was informed by the planning and the logistic capacity of our project. The selection of

participants was based on those experts who responded to an e-mail invitation. We did not specifically select medical doctors trained in geriatric medicine. For our strategies, we used author lists from publications and abstracts and special interest groups focusing on geriatric co-management. However, it is very likely that geriatricians are more interested in geriatric co-management and therefore more likely to respond to an invitation. Unfortunately, the response rate of non-medical professionals was lower than expected. The limitations in the sample size/recruitment are acknowledged in the discussion section.

Reviewer: 2 (Victoria Tang)
1) Table 1: In regards to the Age based characteristics, I'm assuming the categories are not mutually exclusive. Is that correct? The category of "Age <65 years" confused me on first pass.
- Response: In Table 1 the aged based characteristics are reported as mutually exclusive categories. We have added an explanatory note: “The category Age < 65 years refers to studies recruiting patients aged 26 years or older (n = 1), 50 years or older (n = 3), 55 years or older (n = 1), 60 years or older (n = 5).”

2) Some definitions would be helpful, (i.e., team meetings" and "transitional care")
- Response: We thank the reviewer for this suggestion. We have added the following definitions as a footnote to table 2:
2 Team meetings were defined as “case conferences or multidisciplinary meetings in which the geriatrician or geriatrics team interacts with the primary treating physician or other ward staff (e.g. registered nurses, physical therapists) to discuss patients included in the co-management programme”.
3 Medical review was defined as “the prevention of iatrogenic complications through assessment and delivery of interventions that addresses actual or potential problems identified in the assessment”.
4 Rehabilitation was defined as “assessing the need for physical therapy and providing physical and occupational therapy to prevent or reverse functional decline”.
5 Transitional care was defined as “a set of actions designed to ensure the coordination and continuity of health care as patients transfer between different locations or different levels of care in the same location”.

3) Page 14, line 8-10: "11 new indicators were added to the questionnaire": I'd like more information on how and where were they developed from if they were not developed from the systematic review.
- Response: The new indicators were based on the personal experience and knowledge of the participants. This is explained in the methods section: “Participants could suggest additional indicators based on their experience and knowledge.” We added a sentence in the results section to make this clearer: “… four indicators were removed and eleven new indicators were added to the questionnaire. These new indicators were suggested by the Delphi participants.” We have added a new table in the supplement (table 3) defining the new indicators.

Reviewer 3 (Alexander Joeris)
1) The systematic literature review should be described in more details; additionally to the statement which databases were searched, the authors should either explain in the main manuscript or attach as a supplement in more details which search terms and combinations of search terms were used. Additionally, it would be of importance for the reader to explain in more detail how the 12'794 initially identified records were condensed down to 44 text articles for data extraction.
- Response A: We thank the reviewer for this suggestion. We did not include all the details from the review because this has already been reported in our review protocol which is available for free in the Prospero database (reference 12). A detailed search strategy has also been reported in our meta-analysis (reference 5). In order to prevent the publication of duplicate information, we decided not to add more information in the manuscript. Instead, we made a more explicit reference to the available

sources: “The study methodology and search strategy has been detailed elsewhere and is available in a review protocol in the PROSPERO database (CRD42015026033). [5][12]”.
- Response B: We have added the following text at the beginning of the results section: “A total of 12794 titles and abstracts were independently screened by two authors. A total of 335 full text articles were independently assessed for eligibility by two authors. A final 44 manuscripts were included for data extraction. Studies were excluded because they did not report the evaluation of an in-hospital co-management program (n = 248), were an abstract (n = 66), letter to the editor (n = 6) or published in another language (n = 3).”

2) The authors are asked to explain in more detail, why only experts from Europe and North America were selected. Although this fact is mentioned in the methodological considerations of the project and it is recommended to test the indicators in other regions, no rational is given why this was decided initially.
- Response: The choice for including experts from Europe and North America was informed by A) the results from our systematic review, and B) the extent of our network and logistic capacity to recruit participants. From the results of the systematic review, which included manuscripts in five languages, we learned that around 78% of co-management studies were performed in either Europe or North America. For both regions, we had several contacts available to facilitate the recruitment and safeguard a certain level of expertise in the participants.
We have changed the following text in the methodological considerations to explain why we only recruited from two regions: “Thirdly, because the majority of evidence on geriatric co-management originates from North America and Europe, the results of this study may only be valid for these regions. Furthermore, it should be noted that despite the differences between countries in organising their health systems, there were only minimal differences in appropriateness between regions. Validation of the indicators in other countries is recommended.”

3) Additionally I want to encourage the authors to give a reasoning why not more experts were invited for the Delphi process.
- Response: Please see our response to question 3 of reviewer 1. To the best of our knowledge there are no guidelines to determine a minimum sample size for a Delphi study (Keeney et al. J Adv Nurs 2006;53(2):205-212). The number of experts that we aimed to recruit was informed by the planning and the logistic capacity of our project. The selection of participants was based on those experts who responded to an e-mail invitation. While we acknowledge that more participants lead to more data and more information, this does not mean that the results would be different or that the current indicators are not valid.

4) Although the authors highlight the fact that the majority of experts were medical doctors and only a few nurses and one manager, even within the medical doctors one orthopedic surgeon was included only. Looking at the reference list, there would have been a considerably number of orthopedic surgeons which could have participated potentially. As the acceptance of a consensus largely depends on the composition of the Delphi panel, the authors should give a reasoning for the decision to assemble the panel as it was. This especially, as in 87% of the articles the patient population of interest was "surgical".
- Response: Please see our response to question 3 of reviewer 1. We have added the following text to the methodological considerations: “The selection of participants was based on those experts who responded to an e-mail invitation. We did not specifically select medical doctors trained in geriatric medicine. For our strategies, we used author lists from publications and abstracts and special interest groups focusing on geriatric co-management. However, it is very likely that geriatricians are more interested in geriatric co-management and therefore more likely to respond to an invitation.” We do not believe that the small number of non-medical professionals invalidates our results.

5) Delphi study: it is not clear for the reader how the indicators in round 1 were compiled/re-worded/removed or added. It would be reasonable that the authors add a table as supplement which contain the indicators for round 1, which indicators were removed and which added for round 2
- Response: We thank the reviewer for this suggestion. We have added two new tables in the supplement as requested. Please see Supplementary table S2: Indicators removed after round 1 and Supplementary table S3: Indicators added after round 1.

6) Result section: Outcome indicators (p19 line 6): in fig 2 it is written that 17 outcome indicators are considered appropriate and feasible, in the text the authors state 16. Is this a pure reporting mistake or does the figure has to be interpreted differently? If so, please explain better, as otherwise the numbers are misleading.
- Response: Both the numbers in the text and figure are correct. There are 16 outcome indicators who are both appropriate and feasible. There are 17 outcome indicators who are feasible. Therefore, one outcome indicator was considered feasible to measure, but not appropriate.
We have added the following text in the legend of figure 2 to make this clearer: “Note: Of the seventeen outcome indicators who were considered feasible, sixteen were also considered appropriate.”

7) Discussion: p21, line 54ff: it is not surprisingly, but concerning that post-discharge follow-up outcomes were generally not considered feasible by the experts. This indicates the urgent need and importance to develop and implement post-discharge outcome measures which are thought to be feasible.
- Response: We thank the reviewer for this suggestion. We have included this as an explicit recommendation for future research in our conclusion: “Future research should focus on the development of post-discharge outcomes who are feasible to measure …”

8) Methodological considerations: The authors state that the quality of the primary studies was generally poor. Is this a subjective statement or were measures taken to support this statement? If so, please describe/provide more information
- Response: This observation was based on a recently published meta-analysis by our team. We have changed the text in the discussion section to: ”a recent meta-analysis on geriatric co-management programmes observed a high risk of bias and poor reporting of study methodology in published manuscripts.[5]”