Reviewer’s report

Title: Collective intelligence in medical decision-making: a systematic scoping review

Version: 0 Date: 11 Feb 2019

Reviewer: Ralf H. J. M. Kurvers

Reviewer's report:

A review on collective intelligence in medical diagnostics is no easy feat given the heterogeneity of the field, the often ill-defined terms, the existence of many related concepts, and the different research communities working on this topic. I think this manuscript can be helpful in clarifying the concept of collective intelligence in medical diagnostics, and provide some structure to the field, though I feel it needs to be revised to achieve this goal. In its current form, I think it is not yet clear enough, especially regarding (i) the main aim of the paper, (ii) the definitions of the terms used (especially collective intelligence is ill defined), (iii) the content of the studies it is summarizing, and (iv) the inclusion criteria used.

Since there were no page numbers, I used the page numbers of the PDF I received. Please note that the Appendix was missing.

Abstract

P3, L4/5. "Collective intelligence..., refers to the insights of groups".

This is vague. I would give a more intuitive description of what collective intelligence entails.

P3, L10/11. "characterize the current state of research with respect to collective intelligence and medical decision-making”.

This goal is not very concrete. Is it not possible to have a more concrete goal in mind with this scoping review?

P3, L20/21. This sentence seems to be incorrect.

P3, L32/33. "Initial decision task (group vs. individual)".

Without any explanation, it is unclear what is meant with this. It would be helpful to either use clearer terminology, or explain this.

P3, L36/37. "Availability of collective intelligence to participants".
Again, please clarify. According to the definition given earlier in the abstract, this means: "availability of the insights of groups to participants". But what does this mean?

P3, L37 "a better understanding of collective intelligence and its goals may improve medical decision-making".

The goal of collective intelligence? What do the authors have in mind with this statement?

P3, L42/43. "promising tool". I am not sure how the preceding text in the abstract justify the conclusion that collective intelligence is a promising tool. (I do not disagree with the statement, but I do not see how it logically follows from the rest of the abstract.)

Background

P4 When introducing the concept of collective intelligence, a number of classical studies on collective intelligence are not cited. Please add general studies on this topic. (e.g. Krause et al. 2010 TREE Swarm Intelligence in Animals and Humans; Surowiecki 2004 The Wisdom of Crowds, etc).

P4, L6/7 "Groups using collective intelligence".

This is ambiguous. What is meant with this? Can there also be groups not using collective intelligence? In general, I think the word collective intelligence is used a bit too broadly throughout the manuscript, and it would help to use this specific word less often, and be more concrete in what it meant in several places in the manuscript when using this word. This, for example, is also true in the abstract when stating "Availability of collective intelligence to participants", the meaning of which is unclear.

P4, L6/7 why is reference ‘2’ not added to references 3/4 at the end of the sentence?

P4, L9/10 Groups can also be comprised of a combination of experts and novices.

P4, L37. "and its implications for expert medical decision-making lack clarity in the literature"

Please give a few concrete examples to clarify what is meant with this rather vague statement.

P4, L46 "inter-expert collaboration".

Collaboration is usually seen as individuals working directly together (whereas the aggregation of independent decisions is usually not considered collaboration; the authors use a similar distinction on P. 6, L. 12). Therefore, I would not state here that the review focuses on "collaboration" only.
Generally, the background section is relatively short and could be improved by discussing collective intelligence in a broader fashion. It is currently very narrow in scope.

Methods

"did not search for specific group activities that may constitute collective intelligence".

Please clarify what this means. I also do not understand why "specific examples" of collective intelligence were not directly searched for, given that the authors state, "our search strategy aimed to capture any type of collaborative decision-making"

The Appendix was not available to me for reviewing.

This implies that grey literature does not include research on clinical settings. But is this really true?

Please report the values of the inter-rater reliability score, so it is possible to evaluate the reliability of the search. (Rather than reporting that this was done).

"we included studies that applied a collective intelligence to any aspect". Applying a 'collective intelligence' sounds ambiguous. Please clarify this sentence.

When reading the section on data extraction, I was surprised to read that the number and type of raters were not extracted. But then in Table 1, I learn this has been done. Please clarify.

I would not use the heading "results" given that six lines later there is the main section "Results".

In the methods section it would be helpful to have a better overview of the inclusion/exclusion criteria. Perhaps even have a list with all the criteria. Criteria are now often not clearly defined (e.g., "consideration of complex, non-binary data"; but how exactly was the decision made whether a study was "complex" or not?). Moreover, new criteria are being introduced in the results section (e.g., "no studies were included that included non-medical experts"). It is thus not easy to follow the literature search process.

Results

"into a single collective intelligence".
I think it would be better to simply talk about a collective decision or a collective outcome. I would not use the phrasing "collective intelligence" here.

P8, L.44 "expert discussion". This sounds very similar as the "open discussion among group members as discussed in the section on "initial decision task". Perhaps clarify that (if I understand this correctly) how these two discussion types are different (or happening in different phases of the process).

P8, L51 "In four of six individual-based studies, technology facilitated the aggregation of individual opinions into a collective intelligence 17,22,26,27"

When looking at some of these studies, I am not certain if this statement is correct. Take reference 27. This study compares the accuracy of single physicians versus computers. I do not see how this is related to collective intelligence; nor how this is an example of studies in which technology facilitated the aggregation of individual opinions into a collective intelligence.

P9, L35 "Availability of collective intelligence output"

Please explain what is meant with this? I think it would help to reformulate this, because it is unclear what it means to have collective intelligence 'available'. Perhaps collective outcome? Or collective decision?

P10, L.3 "but did not necessarily investigate the accuracy of collective intelligence or its benefit over individual decision-making."

It would be helpful to already state in the introduction that this scoping review is not only focusing on collective accuracy as compared to individual accuracy. In almost all studies that talk about collective intelligence, accuracy is a central concept, so I was surprised that this was not a hard selection criterion. I would thus already in the introduction sketch this broad approach to clarify the scope better to the reader.

P10, L.8. "For example, information needs among participants at a case conference, a traditional means of generating collective intelligence, were resolved through group processes of collective intelligence".

Were resolved through group processes of collective intelligence. Please clarify what this means, as this is hard to comprehend.

P10, L23 "Finally, one study demonstrated that a collaborative technology platform could enhance a group's collective intelligence".

Please explain what this means. Better diagnostic accuracy? Again, the word collective intelligence is used so broadly throughout the text that it is really hard to understand its meaning.
P10, L40/42. When looking at references 20 and 22 I fail to see how they are connected to the concept of collective intelligence. These studies investigate how different groups of raters have different use of ranking. They do not (as far as I can see) relate to anything related to group decision making. Also when looking at the statements in the text I fail to see how this is related to collective intelligence: "finding that specialists varied systematically in the diagnoses they reached, and that experts were more accurate than medical students". I get a bit worried about the suitability of these studies to be included in a review about collective intelligence in medical diagnostics.

P10, L58. "As compared to novices, expert physicians had better diagnostic accuracy and faster decision times."

Again, how is this statement relevant to collective intelligence?

P11, L1 + L30. Why repeat this statement? (In addition, it is also referenced on P10, L60).

Generally, in the results section, I found several claims that, in my opinion, are not backed up by the specific literature references provided. Moreover, when looking at several of these references I do not see an obvious link to collective intelligence. It is possible that the authors have a radically different view of what collective intelligence entails, but then it would be helpful if this view is presented early on so that it becomes clear which aspects are being considered a part of collective intelligence. I think the results section needs to be clarified, and the claims need to follow the cited studies more rigorously and, at places, more concretely.

Discussion

P11, L45 "efficiently resolve uncertainty". What is mean with resolving uncertainty? Did this also resulted in increased decision accuracy? Please clarify such statements. This is one example of ambiguous writing present in several parts of the manuscript.

P11, L48 Broken sentence.

A bit more forward looking into what are important questions to answer in the context of collective intelligence and medical diagnostics

Potential literature to include (depending on exact inclusion criteria):

Hernández-Chan et al. 2012.

Knowledge Acquisition for Medical Diagnosis Using Collective Intelligence.

Journal of Medical Systems. https://doi.org/10.1007/s10916-012-9886-3
Collective intelligence in medical diagnosis systems: A case study.

Medical Decision Making. https://doi.org/10.1177/0272989X17696998

Kurvers et al. 2018. Combining independent decisions increases diagnostic accuracy of reading lumbosacral radiographs and magnetic resonance imaging.
PLOS ONE. https://doi.org/10.1371/journal.pone.0194128

Signed,
Ralf Kurvers

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

Unable to assess

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Yes

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

No
Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

Not relevant to this manuscript

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?
4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?
5. Do you have any other financial competing interests?
6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal