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)	Future Health Today: Co-design of an electronic chronic disease quality improvement tool for use in general practice using a service design approach
AUTHORS	Hunter, Barbara; Biezen, Ruby; Alexander, Karyn; Lumsden, Natalie; Hallinan, Christine; Wood, Anna; McMorrow, Rita; Jones, Julia; Nelson, Craig; Manski-Nankervis, Jo-Anne

VERSION 1 – REVIEW

REVIEWER	Renee Lyons University of Toronto Canada
REVIEW RETURNED	17-Jul-2020

GENERAL COMMENTS	Important topic with methods clearly laid out. Great potential. The paper begins with a discussion of the patients most at risk and with multiple health issues yet the study does not address this population. There is no review of literature on current approaches to tool like this and lessons learned. There is much that has been done on this topic and the participants who were engaged could have benefitted also by being introduced to existing systems such a Kaiser Permanente's and others. The actual tool itself is described but an Appendix giving a snapshot of where the researchers started and ended up would have been useful. The value add of this work is not clear, in the context of existing research and tool development. Millions have been spent on systems - clearly articulate learnings from this work that add to the literature. A more sophisticated analysis of the researchers co-design process would have been useful. There is no critique of the co-design processes used, as well as no comparison across populations regarding agreements and disagreements; e.g., conflicts between patients and practitioners; characteristics of settings that support the use of this tool, conditions in practices that would facilitate its use. Are any theoretical ideas generated from this work? What is the longer term plan for development and testing or knowledge translation? Is there support from government to further develop the tool? Any thoughts on the criteria for assessing quality improvement of the items included?

REVIEWER	Dirk Ramaekers Leuven Institute for Healthcare Policy, Belgium

REVIEW RETURNED	16-Aug-2020

GENERAL COMMENTS	This paper describes the outcomes of the development process of an electronic chronic disease quality improvement tool in a primary care setting. 'Co-design' was used to improve physician engagement with a specific quality improvement program, Future Health Today FHT. A number of questions emerge by reviewing the methods, results and discussion section. - Patients were recruited at the beginning of the project. How were they recruited? Criteria? How were they trained? In the results section only 5 patients appeared to be included, with four over 60. This very small sample is unlikely to be representative. It is unclear from the methods how patient choices in the care process itself were explored. It appears that the focus was placed on solely their participation (what type of participation?) in the recall visits. - Several methods were combined. Some are derived from qualitative research in behavioral sciences. Others are more frequently used by management consultants (e.g. blue sky thinking) or human sciences. No evidence-based approach was used. This would have provided research on similar, already (partially) implemented, systems (in e.g. the US, Finland, Belgium,…). Experiences in other countries should also have been a topic in the discussion section. - Co-design methodology. Isn't this basically an informal consensus methodology? How were psychological influences from dominant participants in e.g. the few controlled for and thus bias avoided? What was exactly the content of the semi-structured interviews and how was this validated? The authors give the impression a limited number of participants (that was unstable over time) was able to give valid results in the development process. - No real world data (of variation in current practices quality e.g. from local audit and feedback data) were used as a starting point to prioritize elements in future services (or was not mentioned?). Story telling reflects an idealized reality, risking a larger implementation gap. - No technological details are provided on the platform and its components nor integration in existing systems. Relationships and possible sponsorships with health or tech-industries should be mentioned. To summarize, this manuscript provides interesting information on the development of an electronic QI tool emphasizing the involvement of practitioners in that process. To a lesser extent, patients' opinions were involved. Worldwide, several similar QI-instruments have been developed, but only a limited number has been implemented in daily practice and has shown to be effective and safe. The authors state that the platform is currently being piloted in a general practice clinic. The manuscript does not state that a call to 40 general practices to be part of a trial of their practice program has been already launched (see website of FHT). From a scientific point of view, especially the results of that trial will be more important provided relevant outcomes and (patient) experience are assessed. Taken into account the these concerns, this manuscript would need major revision. It is possibly more suited for another journal in the domain of medical informatics or qualitative implementation research.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1	Reviewer 2	Changes made
Background/discussion		
The paper begins with a discussion of the patients most at risk and with multiple health issues yet the study does not address this population.		Chronic kidney disease, one of the conditions discussed in the beginning of the paper, was used as the exemplar to assist participants to conceptualise how FHT might work, as is noted throughout the paper. We have further noted that the platform is designed to capture chronic disease more broadly and is currently being expanded beyond CKD to address these multiple conditions (p14, final paragraph of discussion).
There is no review of literature on current approaches to tool like this and lessons learned. There is much that has been done on this topic and the participants who were engaged could have benefitted also by being introduced to existing systems such a Kaiser Permanente's and others.		We have added further background on similar primary care focused QI audit and feedback systems in other jurisdictions and challenges associated with implementation (p3, paragraph 1).
The actual tool itself is described but an Appendix giving a snapshot of where the researchers started and ended up would have been useful.	No technological details are provided on the platform and its components nor integration in existing systems.	We have added an appendix with a snapshot of initial design through to prototype product, which provides a summary of the platform and its components. The intention of the piece is to describe the design process, not the technological specifications.
	Relationships and possible sponsorships with health or tech-industries should be mentioned.	We are not involved in any sponsorship arrangements with health or tech industries, however we are fee-paying members of the Best Practice Partner program which allows

Reviewer 1	Reviewer 2	Changes made
		FHT to integrate with that product.
The value add of this work is not clear, in the context of existing research and tool development. Millions have been spent on systems - clearly articulate learnings from this work that add to the literature.		The aim of this paper was to describe the process of co-designing a tool that meets the needs of Australian general practice – where there is currently a gap that no tool adequately fills. The project tested assumptions from the literature about what works with the end-users, and to explore other factors or issues that they felt were important, based on their experience and clinical expertise, to design a tool that they believe will work for them.
Methods		
	Patients were recruited at the beginning of the project. How were they recruited? Criteria? How were they trained? In the results section only 5 patients appeared to be included, with four over 60. This very small sample is unlikely to be representative. It is unclear from the methods how patient choices in the care process itself were explored. It appears that the focus was placed on solely their participation (what type of participation?) in the recall visits.	The methods section identifies how patients were recruited (p4). We have added inclusion criteria. Patients did not receive training to participate in the co-design session, as their input from their current experience and perspective was sought, with particular focus on their experience of being recalled by the health care professional to general practice for management of chronic health conditions. No further training was required.
	Several methods were combined. Some are derived from qualitative research in behavioral sciences. Others are more frequently used by management consultants (e.g. blue sky thinking) or human sciences. No evidence-based approach was used. This would have provided research	The application of service design methodology, with a co-design component, is evidence based. This has been included in the method section (p3).

Reviewer 1	Reviewer 2	Changes made
	on similar, already (partially) implemented, systems (in e.g. the US, Finland, Belgium,…). Experiences in other countries should also have been a topic in the discussion section.	Relevant previous work in this area has been included in the background section.
A more sophisticated analysis of the researchers co-design process would have been useful. There is no critique of the co-design processes used, as well as no comparison across populations regarding agreements and disagreements; e.g., conflicts between patients and practitioners; characteristics of settings that support the use of this tool, conditions in practices that would facilitate its use.	Co-design methodology. Isn't this basically an informal consensus methodology? How were psychological influences from dominant participants in e.g. the few controlled for and thus bias avoided? What was exactly the content of the semi-structured interviews and how was this validated? The authors give the impression a limited number of participants (that was unstable over time) was able to give valid results in the development process.	The method is explained further in the methods section, addressing these concerns (p3). The intention of the service design approach, and co-design, is to gather a range of perspectives, in the ongoing development of the tool, not consensus. There were few conflicts across populations, but page 13, paragraph 1, identifies the single point of difference – the use of graphs and visual display for patients. Page 12, paragraph 2, provides contexts/conditions participants felt were more suited to use of the tool. Further piloting will identify the most effective implementation settings/conditions/etc. The schedules used in the co-design sessions have been attached as an appendix.
	No real world data (of variation in current practices quality e.g. from local audit and feedback data) were used as a starting point to prioritize elements in future services (or was not mentioned?). Story telling reflects an idealized reality, risking a larger implementation gap.	Thank you for your comment. The focus of the project was to develop a system that enabled audit and feedback. Participants had familiarity with the systems currently available to them, and the shortcomings of these. The intention was not to fix those systems, rather to develop a system that worked for them to enact QI activities, not just provide a report.
Conclusions		

Reviewer 1	Reviewer 2	Changes made
What is the longer term plan for development and testing or knowledge translation? Is there support from government to further develop the tool?	To summarize, this manuscript provides interesting information on the development of an electronic QI tool emphasizing the involvement of practitioners in that process. To a lesser extent, patients' opinions were involved. Worldwide, several similar QI-instruments have been developed, but only a limited number has been implemented in daily practice and has shown to be effective and safe. The authors state that the platform is currently being piloted in a general practice clinic. The manuscript does not state that a call to 40 general practices to be part of a trial of their practice program has been already launched (see website of FHT). From a scientific point of view, especially the results of that trial will be more important provided relevant outcomes and (patient) experience are assessed. Taken into account the these concerns, this manuscript would need major revision. It is possibly more suited for another journal in the domain of medical informatics or qualitative implementation research.	Thank you for your comment and opportunity to update the status of the 'next steps'. We have provided a statement at the end of the discussion that outlines the next steps for the FHT project, namely piloting in two general practices was completed early 2020, further optimisation is occurring in 2020-21, and a pragmatic cluster RCT will commence in late 2021 (advertisement of the trial commenced in late August, well after the paper was submitted). Two advisory groups are being established to guide the next steps of the project, one comprising 12-15 health practitioners (GPs, nurses, practice managers) and the other comprising 12-15 patient/carers.
Any thoughts on the criteria for assessing quality improvement of the items included?		We intend to use the CP-FIT framework, as described by Brown et al, to examine the effectiveness of the FHT platform in improving QI. This is included in the discussion (p14).
Are any theoretical ideas generated from this work?		The process seemed to be an effective one, however any contribution to broader theoretical discussion must wait until we see the fruits of

Reviewer 1	Reviewer 2	Changes made
		our labours – i.e., after piloting, evaluating and refining.
Formatting		
1. No corresponding author's email address in main doc: - Please provide a "Corresponding author's email address" in your main document file as shown in ScholarOne.		Apologies, now provided.
2. Please revise your title so that it includes your study design. This is the preferred format for the journal.		Apologies, we have adjusted.
3. Please revise the abstract so that it is following the structured abstract recommended in BMJ Open's Instructions for Authors for research articles. See: https://protect-au.mimecast.com/s/YUA2C6X1Pysr1J8GPt5-VU9?domain=bmjopen.bmj.com You do not have to include all methods sections if these are not applicable to your study.		Apologies, revised to required format.

VERSION 2 – REVIEW

REVIEWER	Renee Lyons University of Toronto and Dalhousie University Canada
REVIEW RETURNED	21-Oct-2020

GENERAL COMMENTS	This is the second time I have reviewed this paper. The authors have addressed concerns. The areas that perhaps need discussion related to on-clinical factors in chronic disease management such as SES (poverty/homelessness); mental health and or cultural factors in QI and followup; as well as multi-morbidities. Is there a place for QI and dashboards to include such issues? Were they raised by respondents? If not why, not?

VERSION 2 – AUTHOR RESPONSE

Reviewer 1:	We have inserted 2 paragraphs into the results section addressing these questions. In summary, the respondents did raise these

This is the second time I have reviewed this paper. The authors have addressed concerns. The areas that perhaps need discussion related to on-clinical factors in chronic disease management such as SES (poverty/homelessness); mental health and or cultural factors in QI and followup; as well as multi-morbidities. Is there a place for QI and dashboards to include such issues? Were they raised by respondents? If not why, not?	issues and were unable to identify a clear technological solution to adequately incorporate non-clinical factors that are not captured in the electronic medical record system. These issues required additional consideration to determine which features could be embedded in the FHT technology, and which would form part of the broader implementation of FHT within a quality improvement framework. Similarly, issues relating to the presentation of multiple conditions/multi-morbidity on the platform did not resolve with a single solution, and indicated to the research and development team that a multi-pronged approach may be required. Development was to focus on the prototype with a single condition to test if the concept was both possible and useful.