

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)	The effectiveness of toolkits as knowledge translation strategies for integrating evidence into clinical care: A systematic review
AUTHORS	Yamada, Janet; Shorkey, Allyson; Barwick, Melanie; Widger, Kimberley; Stevens, Bonnie

VERSION 1 - REVIEW

REVIEWER	Nasreen Jessani South Africa
REVIEW RETURNED	05-Dec-2014

GENERAL COMMENTS	<p>The authors have done a good job of embedding this study within the greater literature (except for some references suggested in the comments). Below are a few thoughts on how the study could be strengthened with the hopes that it is helpful for the authors.</p> <p>The study oscillates between talking about KT toolkits and toolkits as one form of KT efforts. As one reads the entire study it becomes more clear that the authors intend to refer to a study of toolkits in general (multiple types are mentioned: Literacy and Numeracy Education toolkits, Fall prevention toolkits etc) as an effective form of KT. It therefore appears that several statements in the study need rephrasing and the title needs to better reflect the objectives of the study. This is important given that KT and KM toolkits do exist within the health sector as well as the development sector in particular. For this reason, given that the toolkits referred to are not definitive KT toolkits, it would seem inaccurate to represent them as so.</p> <p>It is unclear whether the authors consider the toolkits and guidelines as KT efforts or whether the authors of the original studies refer to them as KT efforts. This needs to be made clearer in order to avoid misrepresentation. It seems fair to refer to toolkits and guidelines as inputs or interventions to affect behavior change but it the same would not occur in the context of KT.</p> <p>The authors cite CIHR for a description of KT but fails to include the full text that highlights that the KT process “takes place within a complex system of interactions between researchers and knowledge users which may vary in intensity, complexity and level of engagement depending on the nature of the research and the findings as well as the needs of the particular knowledge user.” The concept of KT therefore is often used in contexts of research knowledge in particular. The participants in the studies included are clinicians, caregivers, or patients that are involved in implementing the guidelines in the toolkits. They are not researchers who are using toolkits and guidelines as KT strategy for knowledge users. They are the knowledge users. For this reason, the use of the term</p>
-------------------------	---

	<p>KT within this study is skewed and doesn't quite present KT in the form intended and in the form understood by the KT community. On this note, it would seem that the introduction focuses on KT as a concept and refers to studies that look at KT as a research to policy and practice endeavor. This study however focuses on toolkits as one form of KT to guide evidence informed practices/behaviors. There is a disjoint therefore between what is being presented in the introduction and the objectives of the study. This circles back perhaps to the understanding of KT.</p> <p>The authors reference Grimshaw et al as reference No [3] but the actual article from which the results the authors cite come from another article by Grimshaw JM, et al: Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technol Assess 2004, 8:iii-72. Furthermore, the statement about Prior et al [7] findings seem to misrepresent what Prior et al actually state which is: "effective implementation strategies included multifaceted interventions, interactive education and clinical reminder systems" The authors of this study state that the results indicate that "Prior et al reported multifaceted interventions such as interactive education sessions were more effective." This could unintentionally mislead the readers. Authors should double check citations as well as original author intentions when citing studies. It would seem that the studies come from various countries and therefore contribute towards a sense of generalizability. It would be interesting to include the country(ies) of the study in Table 1 as well as in the discussions.</p> <p>Also it might be interesting to separate out any nuances of effectiveness on behavior change among health care professionals versus that of patients versus that of caregivers in the various studies. It might say more about the audience of the strategy rather than just the medium of the strategy and clarify the conclusions. It might also help determine whether the 8 studies really are sufficient to draw convincing conclusions given the diversity of target audiences and contexts within the 8 studies.</p> <p>The study refers to single KT interventions and multifaceted KT interventions. It is not clear what each of these are but it would seem that the use of toolkits are single interventions that could be part of a multifaceted KT strategy. Furthermore a single intervention could be used multiple times and multifaceted interventions at one single time. Without understanding what the authors mean but these terms it is difficult for the reader to comprehend the intent. A better definition (and example) of single vs multifaceted would be helpful.</p> <p>Oftentimes the authors use KT when it seems like they might mean "communication" or "knowledge dissemination" which is not KT in the way that CIHR and other KT scholars intend for it to be used. Eg pg 12: "Mediums used to deliver KT strategies include informational sheets, posters, pockets guides, educational modules etc"</p> <p>The methods section doesn't mention "KT" as a search term so can the authors be certain that guidelines pertaining to KT (versus guidelines/toolkits being a form of KT) have been captured? It is understandable that the authors' search occurred earlier in the year but it may be worth checking if this recent publication fits the selection criteria and adds to the review: Implementation of Resources to Support Patient Physical Activity Through Diabetes Centres in Atlantic Canada: The Effectiveness of Toolkit-Based Physical Activity Counselling</p>
--	--

	<p>http://www.canadianjournalofdiabetes.com/article/S1499-2671%2814%2900086-0/abstract?cc=y Another article that might be important to reference in the introduction is the suggestion for evidence based toolkits as a way to influence evidence based behavior: Implementing Evidence-Based Practices for Persons With Severe Mental Illnesses (http://ps.psychiatryonline.org/doi/abs/10.1176/appi.ps.52.1.45)</p> <p>The study selection mentioned criteria for KT goals of the studies selected but not necessarily KT goals of the toolkits. The authors could perhaps assist in reconciling this for the readers. Furthermore, were studies that were more qualitative in nature excluded from the review? Mention of the EHPP tool suggests that only studies evaluating toolkits quantitatively were included but being more explicit about this (and the limitations) would be helpful.</p> <p>In the data extraction section authors mention that “evidence underpinning development.....choice of KT strategies was extracted.” Were these KT strategies of the toolkits or KT strategies of the studies included in the review?</p> <p>The conclusions of the study are valid in the sense that there is a need for evidence to inform the guidelines. However sometimes toolkits are not intended to be prescriptive but rather more as guides that can be adapted or applied depending on the context. Perhaps the interchangeable use of the word toolkit as well as guideline is what causes some of the opacity in methods to determine their effectiveness. Addressing the points above may help in strengthening the assertions in the conclusions. As they currently stand, they need further support. For instance, “a planful approach” is unclear. Where KT goals are mentioned (“a planful approach...for achieving intended KT goals”), it seems the authors may be referring to behaviour change goals. Referring to toolkits as a multifaceted intervention (when perhaps they are one aspect of a multifaceted intervention) unless authors are referring to a toolkit that provides multiple interventions within it. KT strategies keep being nested as educational materials, audits and feedback. This seems rather narrow.</p> <p>In the discussion, authors refer to barriers and facilitators to the toolkits development which is perplexing as most of the paper talks about toolkit use not toolkit development. All this needs more clarity. Furthermore, the summary needs to emphasize that the toolkits can be a promising strategy...for clinical use (since it is only within this context that this study is embedded). This will alleviate challenge from those using toolkits in other contexts. However given that there are only 8 eligible studies of which each deal with different users (patients vs caretakers vs clinicians) it would need a lot of convincing to present definitive conclusions.</p> <p>The title would need rephrasing once the clarifications above have been addressed.</p>
--	---

REVIEWER	Alison Kitson University of Adelaide South Australia Australia
REVIEW RETURNED	11-Dec-2014

GENERAL COMMENTS	<p>well written and clear paper.</p> <p>approach taken for systematic review seems reasonable. may need to explain more clearly how you synthesised your findings. You mentioned Proctor's implementation outcomes but these seven dimensions don't seem to appear in the discussion</p> <p>would have liked to have had a summary table with the range of elements you found in different toolkits, frequency of use and effectiveness</p> <p>would also have liked more in the discussion about how these elements might work together (or not) from a theoretical perspective</p> <p>reference to the Stevens et al study in the discussion was confusing as it was not part of those that were reviewed. Either need to limit reference or draw on other examples</p>
-------------------------	---

REVIEWER	Pierre PLUYE McGill University
REVIEW RETURNED	15-Dec-2014

GENERAL COMMENTS	<p>Overall, this manuscript is well-written, and concerns an important topic (knowledge translation). The present manuscript reports a systematic review of the literature on the effectiveness of KT toolkits. The definition of KT is correct and well-referenced (CIHR, Strauss, Grimshaw). KT strategies include (but are not limited to) educational material. The authors' review of previous reviews suggests multifaceted KT interventions (MKTi) include multiple strategies (education, opinion-leader, audit, feedback, reminders); among them, interactive education sessions may be the most effective, while "no definitive evidence in favour of using either multifaceted or single KT interventions" (Grimshaw & Eccles, 2012).</p> <p>However, this manuscript faces two major issues. First, the notion of 'toolkit' is unclear. Authors stated that "The complexity of integrating evidence into clinical care and ensuring it is effectively implemented can be addressed through the development and evaluation of what is often referred to as KT and implementation toolkits". However, there are no references directly supporting this claim ("often"), and the definition of KT toolkit is not clearly supported by previous work. Such toolkit is defined as "a set of tools or strategies used for the purpose of educating and/or promoting behavior change in health care professionals, patients or caregivers in a hospital setting" for "facilitating the implementation of evidence into clinical care." An example is given: a set of 50 guidelines (Registered Nurses' Association Ontario. Nursing best practice guidelines). Thus, the differences between MKTi vs. educational material vs. KT toolkits are unclear for instance. In fact, authors retained 8 studies, and concluded that "The combination and types of KT strategies embedded within toolkits varied across studies but included predominantly educational materials." This issue does not allow reviewers/readers to understand this manuscript, and its contribution to scientific knowledge. It raises the two following key questions. Q1. What is a KT toolkit? a clear conceptualisation is needed (a clear definition of the concept, a clear conceptual framework or a direct reference to a previous conceptual work, and illustrative examples – e.g., examples of MKTi and KT strategies relevant and representative of all aspects of the concept).</p>
-------------------------	--

	<p>Q2. Then, what is new/different from recent systematic effectiveness reviews of MKTI and reviews of educational material? a clear description of studies that were included vs. studies included in previous reviews vs. studies never included in other reviews is needed.</p> <p>Second, another major issue is that the structure of the manuscript does not rigorously follow the PRISMA statement. Thus, method and result sections can be rewritten with an expert, e.g., a member of the Cochrane collaboration, using standard guidelines for effectiveness systematic reviews.</p> <p>1. Eligibility criteria are not specifically mentioned. Study designs of interest were not mentioned, while the review is about effectiveness (usually only RCT and CCT).</p> <p>Page 7: Authors stated “Studies were retrieved if they met the following criteria: 1) evaluated the effectiveness of a toolkit to support the integration of evidence into clinical care, either alone or embedded within a larger multi-faceted intervention, and if 2) the KT goal(s) of the study were to educate and/or change practice behaviours and/or clinical outcomes in healthcare settings.” This must be presented in terms of eligibility criteria (first method section on page 6). In fact, this sentence confirms the lack of conceptualization and definition of the ‘toolkit’ notion. Eligibility criteria should clearly state what counts as a toolkit, and what does not. In addition, studies were not “retrieved”, but selected or included.</p> <p>2. Sources of information: Main health bibliographic databases were searched. Was a specialized librarian involved? This must be specified.</p> <p>3. Identification of relevant studies: The example of the search strategy from MEDLINE fits with the lack of conceptualisation. The literature was apparently searched only for what researchers called themselves a ‘tool’, ‘toolbox’, or ‘toolkit’. However, many researchers probably studied KT toolkits, while they did not use the word ‘toolkit’ for example.</p> <p>Page 6: Authors stated “The evaluation search terms used in (...) were based on published optimized search strategies.[10-12]” where the references are interesting (PsycINFO search strategies identified methodologically sound therapy studies and review articles for use by clinicians and researchers; Developing optimal search strategies for detecting clinically sound treatment studies in EMBASE; Optimizing search strategies to identify randomized controlled trials in MEDLINE), but cannot replace a specialized librarian considering the standards of effectiveness reviews</p> <p>4. Selection of relevant studies: Two independent reviewers did 10 studies. It is not specified whether all other studies were independently selected by two reviewers, and what was the final crude agreement, and inter-rater reliability score, e.g., Cohen’s kappa.</p> <p>5. Critical appraisal: No risk of bias assessment, while this is encouraged by the Cochrane collaboration (usual standard for effectiveness systematic reviews). Authors stated “The EPHPP tool can be used to evaluate multiple study designs, including RCTs, casecontrols, cohort and cross-sectional studies.”; however, some of these designs can hardly help for an effectiveness review. Thus, this sentence is puzzling. In addition, authors stated “For the purpose of this review, the authors a priori decided to only report on studies with a strong or moderate global rating.” This is discouraged by the Cochrane collaboration. All relevant studies are usually included and results synthesized; specifically, results of the appraisal are of great</p>
--	---

	<p>interest as they can lead to rich sensitivity analysis (comparing lower vs. higher quality studies for instance).</p> <p>6.Data extraction: The following sentence is very confusing: “Because many studies embedded the toolkit into a multi-faceted intervention and did not evaluate the toolkit separately, information regarding all of the components of the study intervention was extracted”. Thus, what was assessed? This seems to confirm the lack of conceptualisation.</p> <p>7.Synthesis: The data analysis must be specified. There is no tabulation of quantitative results, no forest plot. Authors did not mention that a meta-analysis was not feasible. Results seem to report some type of narrative-like synthesis, while there is no clear narrative synthesis method.</p>
--	--

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

The authors have done a good job of embedding this study within the greater literature (except for some references suggested in the comments). Below are a few thoughts on how the study could be strengthened with the hopes that it is helpful for the authors.

1) The study oscillates between talking about KT toolkits and toolkits as one form of KT efforts. As one reads the entire study it becomes more clear that the authors intend to refer to a study of toolkits in general (multiple types are mentioned: Literacy and Numeracy Education toolkits, Fall prevention toolkits etc.) as an effective form of KT. It therefore appears that several statements in the study need rephrasing and the title needs to better reflect the objectives of the study. This is important given that KT and KM toolkits do exist within the health sector as well as the development sector in particular. For this reason, given that the toolkits referred to are not definitive KT toolkits, it would seem inaccurate to represent them as so. It is unclear whether the authors consider the toolkits and guidelines as KT efforts or whether the authors of the original studies refer to them as KT efforts. This needs to be made clearer in order to avoid misrepresentation. It seems fair to refer to toolkits and guidelines as inputs or interventions to affect behaviour change but it the same would not occur in the context of KT.

Response: We have made revisions to the paper to indicate that the review focuses on the study of toolkits in general as an effective form of KT. We consider the toolkits as KT efforts. Toolkits are one form of KT to guide evidence informed practices/behaviours.

2) The authors cite CIHR for a description of KT but fails to include the full text that highlights that the KT process “takes place within a complex system of interactions between researchers and knowledge users which may vary in intensity, complexity and level of engagement depending on the nature of the research and the findings as well as the needs of the particular knowledge user.” The concept of KT therefore is often used in contexts of research knowledge in particular. The participants in the studies included are clinicians, caregivers, or patients that are involved in implementing the guidelines in the toolkits. They are not researchers who are using toolkits and guidelines as KT strategy for knowledge users. They are the knowledge users. For this reason, the use of the term KT within this study is skewed and doesn't quite present KT in the form intended and in the form understood by the KT community.

On this note, it would seem that the introduction focuses on KT as a concept and refers to studies that look at KT as a research to policy and practice endeavor. This study however focuses on toolkits as one form of KT to guide evidence informed practices/behaviors. There is a disjoint therefore between what is being presented in the introduction and the objectives of the study. This circles back perhaps to the understanding of KT.

Response: Unfortunately, we find the reviewer’s rather unclear. From our perspective, the CIHR

definition does include 'dissemination' and we have revised and included the additional definition of KT by CIHR in the introduction on Page 4:

Knowledge translation (KT) is a complex process occurring between researchers and knowledge users that includes the "synthesis, dissemination, exchange and ethically sound application of knowledge to improve health... provide more effective health services and products, and strengthen the health care system." [1] The degree of engagement in the KT process may be influenced by factors such as the research results and needs of the knowledge user. [1]

We agree that KT refers to research knowledge, as this is what is being disseminated in toolkit format. The intended audiences for these toolkits are indeed non-researchers, as stated, but we are unclear as to why this reviewer sees sharing of research knowledge with non-academic knowledge users as not being KT.

3) The authors reference Grimshaw et al as reference No [3] but the actual article from which the results the authors cite come from another article by Grimshaw JM, et al: Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technol Assess* 2004, 8:iii-72. Furthermore, the statement about Prior et al [7] findings seem to misrepresent what Prior et al actually state which is: "effective implementation strategies included multifaceted interventions, interactive education and clinical reminder systems" The authors of this study state that the results indicate that "Prior et al reported multifaceted interventions such as interactive education sessions were more effective." This could unintentionally mislead the readers. Authors should double check citations as well as original author intentions when citing studies.

Response: We have revised the references above to include the Grimshaw 2004 paper and are replacing the Prior paper with an updated reference by Squires et al., 2014 on Page 5:

In a recent systematic review by Squires and colleagues (2014) there was no definitive evidence in favour of using either multifaceted or single KT interventions.[4]

Squires JE, Sullivan K, Eccles MP, Worswick J, Grimshaw JM. Are multifaceted interventions more effective than single-component interventions in changing health-care professionals' behaviours? An overview of systematic reviews. *Implementation science* : IS. 2014;9:152.

4) It would seem that the studies come from various countries and therefore contribute towards a sense of generalizability. It would be interesting to include the country(ies) of the study in Table 1 as well as in the discussions.

Response: We have updated Table 1 and have included the countries of the studies.

5) Also it might be interesting to separate out any nuances of effectiveness on behavior change among health care professionals versus that of patients versus that of caregivers in the various studies. It might say more about the audience of the strategy rather than just the medium of the strategy and clarify the conclusions. It might also help determine whether the 8 studies really are sufficient to draw convincing conclusions given the diversity of target audiences and contexts within the 8 studies.

Response: We have included this information in Table 1 under Toolkit target.

6) The study refers to single KT interventions and multifaceted KT interventions. It is not clear what each of these are but it would seem that the use of toolkits are single interventions that could be part of a multifaceted KT strategy. Furthermore a single intervention could be used multiple times and multifaceted interventions at one single time. Without understanding what the authors mean but these terms it is difficult for the reader to comprehend the intent. A better definition (and example) of single vs multifaceted would be helpful.

Response: On Page 4, we define multifaceted strategies:

Evidence-based KT strategies for linking research evidence and clinical practice include but are not limited to printed educational materials, educational meetings, educational outreach, the use of local opinion leaders, audit and feedback, and reminders.[3] These strategies have been used alone as single interventions or as multifaceted interventions, which consist of two or more strategies or variations of the same strategies (e.g., educational materials) delivered in combination to change practice. Multifaceted KT interventions can also include more than one type of strategy such as education, reminders, and audit and feedback.[5,6]

7) Often times the authors use KT when it seems like they might mean “communication” or “knowledge dissemination” which is not KT in the way that CIHR and other KT scholars intend for it to be used. Eg pg 12: “Mediums used to deliver KT strategies include informational sheets, posters, pockets guides, educational modules etc”

Response: In reviewing the CIHR definition, as above, it is clear the KT does include dissemination. In plain language, we view KT as sharing research evidence in ways that enable knowledge users to understand the evidence. For the purpose of this paper, we define knowledge translation (KT) strategies for linking research evidence and clinical practice to include (but are not limited to) printed educational materials, educational meetings, educational outreach, the use of local opinion leaders, audit and feedback, and reminders (Page 4).

8) The methods section doesn't mention “KT” as a search term so can the authors be certain that guidelines pertaining to KT (versus guidelines/toolkits being a form of KT) have been captured? It is understandable that the authors' search occurred earlier in the year but it may be worth checking if this recent publication fits the selection criteria and adds to the review: Implementation of Resources to Support Patient Physical Activity Through Diabetes Centres in Atlantic Canada: The Effectiveness of Toolkit-Based Physical Activity Counselling

<http://www.canadianjournalofdiabetes.com/article/S1499-2671%2814%2900086-0/abstract?cc=y>

Another article that might be important to reference in the introduction is the suggestion for evidence based toolkits as a way to influence evidence based behavior: Implementing Evidence-Based Practices for Persons With Severe Mental Illnesses

(<http://ps.psychiatryonline.org/doi/abs/10.1176/appi.ps.52.1.45>)

Response: We did not include KT in the search terms as toolkit is the KT strategy of interest, and we were not intending to compare toolkits to other KT strategies. The paper: Implementation of resources to support patient physical activity through diabetes centres in Atlantic Canada: the effectiveness of toolkit-based physical activity counselling is relevant to this review; however, our search strategy was conducted up to November 2013. If we were to update our search, we would include this paper in the review. The paper by Torrey and colleagues: Implementing evidence-based practices for persons with severe mental illnesses was captured in our search. This article provides a rationale for toolkits as a method to influence evidence-based behaviours. We have added reference to this paper in the introduction section on Page 5:

Evidence-based toolkits can be used as a method to influence evidence-based practice behaviours for guideline implementation through the inclusion of exemplar policies, training aids and quality audit materials.[10,11]

9) The study selection mentioned criteria for KT goals of the studies selected but not necessarily KT goals of the toolkits. The authors could perhaps assist in reconciling this for the readers. Furthermore, were studies that were more qualitative in nature excluded from the review? Mention of the EHPP tool suggests that only studies evaluating toolkits quantitatively were included but being more explicit about this (and the limitations) would be helpful.

Response: The study selection mentioned criteria for KT goals of the studies selected but not necessarily KT goals of the toolkits as not all of the studies focused solely on evaluating toolkits. On Page 12, we report that three of the eight studies evaluated the toolkit as a single intervention and five studies embedded the toolkit into a multi-strategy intervention. One of the limitations of the study was that we focused on studies that evaluated the effectiveness of a toolkit to support the integration of evidence into clinical care; therefore, the studies included in this review reported quantitative results. We added this sentence in the Discussion section on Pages 18-19.

10) In the data extraction section authors mention that “evidence underpinning development.....choice of KT strategies was extracted.” Were these KT strategies of the toolkits or KT strategies of the studies included in the review?

Response: The choice of KT strategies refer to the strategies/resources in the toolkits.

11) The conclusions of the study are valid in the sense that there is a need for evidence to inform the guidelines. However sometimes toolkits are not intended to be prescriptive but rather more as guides that can be adapted or applied depending on the context. Perhaps the interchangeable use of the word toolkit as well as guideline is what causes some of the opacity in methods to determine their effectiveness. Addressing the points above may help in strengthening the assertions in the conclusions. As they currently stand, they need further support. For instance, “a planful approach” is unclear. Where KT goals are mentioned (“a planful approach....for achieving intended KT goals”), it seems the authors may be referring to behaviour change goals. Referring to toolkits as a multifaceted intervention (when perhaps they are one aspect of a multifaceted intervention) unless authors are referring to a toolkit that provides multiple interventions within it. KT strategies keep being nested as educational materials, audits and feedback. This seems rather narrow.

Response: Our reference to ‘planful approach for achieving intended KT goals’ refers to the need to identify the KT goal that is being addressed by the toolkit strategy. KT goals can be: to inform, share knowledge, build awareness, change practice, change behaviour, inform policy, or to commercialize an innovation). We have clarified this in the Methods section on Page 7 and in the abstract. KT goal(s) of the study were to inform, share knowledge, build awareness, change practice, change behaviour (in the public), and /or clinical outcomes in health care settings, inform policy, or to commercialize an innovation As for the multifaceted issue, we address this earlier – see item #6.

12) In the discussion, authors refer to barriers and facilitators to the toolkits development which is perplexing as most of the paper talks about toolkit use not toolkit development. All this needs more clarity. Furthermore, the summary needs to emphasize that the toolkits can be a promising strategy...for clinical use (since it is only within this context that this study is embedded). This will alleviate challenge from those using toolkits in other contexts. However given that there are only 8 eligible studies of which each deal with different users (patients vs caretakers vs clinicians) it would need a lot of convincing to present definitive conclusions.

Response: On Page 16, we added a sentence to clarify why we included a discussion on barriers assessment and the practice context:

Further research is needed on how the toolkit was developed, and the influence of the practice context as these factors may influence study outcomes.

13) The title would need rephrasing once the clarifications above have been addressed.

Response: We have revised the title of the review to:

The effectiveness of toolkits as knowledge translation strategies for integrating evidence into clinical care: A Systematic review.

Reviewer: 2

1) -well written and clear paper.

-approach taken for systematic review seems reasonable. may need to explain more clearly how you synthesised your findings. You mentioned Proctor's implementation outcomes but these seven dimensions don't seem to appear in the discussion

Response: We added the following sentence in the Discussion section on Page 16:

Proctor's taxonomy for implementation outcomes were extracted from studies where possible as these outcomes could be used to indicate successful implementation of the toolkit within the health care system.

2) -would have liked to have had a summary table with the range of elements you found in different toolkits, frequency of use and effectiveness

Response: We have provided details about the range of elements found in the different toolkits and their effectiveness in the moderately to strongly rated studies in this review (Table 1). Given the weak methodology of the remaining studies, we did not include them in the table.

3)- would also have liked more in the discussion about how these elements might work together (or not) from a theoretical perspective

Response: This is an important point that requires further study; therefore, we have included this suggestion in the Discussion section on Page 19:

In conclusion, further research is necessary to confirm the effectiveness of toolkits as a KT strategy. Future studies should also include a theoretical approach to assess the implementation effectiveness of the individual components of the toolkits to determine which strategy or combination of strategies contributes to changing and optimizing clinical care.

4) - reference to the Stevens et al study in the discussion was confusing as it was not part of those that were reviewed. Either need to limit reference or draw on other examples

Response: We have removed reference to the studies by Stevens and Yamada in the Discussion section.

Reviewer: 3

Overall, this manuscript is well-written, and concerns an important topic (knowledge translation). The present manuscript reports a systematic review of the literature on the effectiveness of KT toolkits. The definition of KT is correct and well-referenced (CIHR, Strauss, Grimshaw). KT strategies include (but are not limited to) educational material. The authors' review of previous reviews suggests multifaceted KT interventions (MKTi) include multiple strategies (education, opinion-leader, audit, feedback, reminders); among them, interactive education sessions may be the most effective, while "no definitive evidence in favour of using either multifaceted or single KT interventions" (Grimshaw & Eccles, 2012).

However, this manuscript faces two major issues. First, the notion of 'toolkit' is unclear. Authors stated that "The complexity of integrating evidence into clinical care and ensuring it is effectively implemented can be addressed through the development and evaluation of what is often referred to as KT and implementation toolkits". However, there are no references directly supporting this claim ("often"), and the definition of KT toolkit is not clearly supported by previous work. Such toolkit is defined as "a set of tools or strategies used for the purpose of educating and/or promoting behavior change in health care professionals, patients or caregivers in a hospital setting" for "facilitating the implementation of evidence into clinical care." An example is given: a set of 50 guidelines (Registered Nurses' Association Ontario. Nursing best practice guidelines).

Thus, the differences between MKTi vs. educational material vs. KT toolkits are unclear for instance.

In fact, authors retained 8 studies, and concluded that “The combination and types of KT strategies embedded within toolkits varied across studies but included predominantly educational materials.” This issue does not allow reviewers/readers to understand this manuscript, and its contribution to scientific knowledge. It raises the two following key questions.

1) What is a KT toolkit? a clear conceptualisation is needed (a clear definition of the concept, a clear conceptual framework or a direct reference to a previous conceptual work, and illustrative examples – e.g., examples of MKTI and KT strategies relevant and representative of all aspects of the concept).

Response: We have revised the definition of a toolkit based on one of the co-author’s (Barwick) recent scoping review of toolkits (Page 5):

For the purposes of this review, 'toolkit' was defined as packages of multiple resources that codify explicit knowledge (e.g., templates, pocket cards guidelines, algorithms), and are used to educate and/or facilitate behaviour change.[9]

Then, what is new/different from recent systematic effectiveness reviews of MKTI and reviews of educational materials? A clear description of studies that were included vs. studies included in previous reviews vs. studies never included in other reviews is needed.

Response: The strategies in MKTI and educational materials are stand alone resources and are not presented as a toolkits.

2) Second, another major issue is that the structure of the manuscript does not rigorously follow the PRISMA statement. Thus, method and result sections can be rewritten with an expert, e.g., a member of the Cochrane collaboration, using standard guidelines for effectiveness systematic reviews.

Response: We followed the PRISMA statement and have submitted this along with the manuscript. The limitations of this review is that we were not able to conduct a meta-analysis as there was considerable variability in the studies included in the review in terms of the study participants and outcomes. We have added this information in the manuscript on Page 6: (The methods for this review were based on the PRISMA checklist (<http://www.prisma-statement.org/2.1.2%20-%20PRISMA%202009%20Checklist.pdf>)).

3) Eligibility criteria are not specifically mentioned. Study designs of interest were not mentioned, while the review is about effectiveness (usually only RCT and CCT). This must be presented in terms of eligibility criteria (first method section on page 6). In fact, this sentence confirms the lack of conceptualization and definition of the ‘toolkit’ notion. Eligibility criteria should clearly state what counts as a toolkit, and what does not. In addition, studies were not “retrieved”, but selected or included.

Response: On Page 7 we describe the eligibility criteria:

Studies were included if they met the following criteria: 1) evaluated the effectiveness of a toolkit to support the integration of evidence into clinical care, either alone or embedded within a larger multi-strategy intervention, and if 2) the KT goals(s) of the study were to inform, share knowledge, build awareness, change practice, change behaviour (in the public), and /or clinical outcomes in health care settings, inform policy, or to commercialize an innovation. Studies published in languages other than English, thesis dissertations, and studies published in non- peer reviewed journals or in abstract form only were excluded. All study designs were included. Reference lists from included papers were screened for additional studies. Please refer to comment 1 for a definition of a toolkit.

4) Sources of information: Main health bibliographic databases were searched. Was a specialized librarian involved? This must be specified.

Response: A hospital librarian, very experienced in systematic reviews, assisted in developing and

conducting the search strategy. This information is stated on Page 6:

A systematic literature search of 4 electronic databases, MEDLINE (1946-Nov 2013), Embase (1947-Nov 2013), PsycINFO (1806- Nov 2013) and CINAHL (1981-Nov 2013), was conducted by a library information specialist.

5) Identification of relevant studies: The example of the search strategy from MEDLINE fits with the lack of conceptualisation. The literature was apparently searched only for what researchers called themselves a 'tool', 'toolbox', or 'toolkit'. However, many researchers probably studied KT toolkits, while they did not use the word 'toolkit' for example.

Response: We focused our review on the term 'toolkits' as a KT strategy and searched for strategies that referred to toolkits.

6) Authors stated "The evaluation search terms used in (...) were based on published optimized search strategies.[10-12]" where the references are interesting (PsycINFO search strategies identified methodologically sound therapy studies and review articles for use by clinicians and researchers; Developing optimal search strategies for detecting clinically sound treatment studies in EMBASE; Optimizing search strategies to identify randomized controlled trials in MEDLINE), but cannot replace a specialized librarian considering the standards of effectiveness reviews 4. Selection of relevant studies: Two independent reviewers did 10 studies. It is not specified whether all other studies were independently selected by two reviewers, and what was the final crude agreement, and inter-rater reliability score, e.g., Cohen's kappa.

Response- First, all titles and abstracts were screened independently by two reviewers (WL and TH). We calculated a percent agreement for the two individuals who screened the articles for relevance. For the two individuals who rated the papers for methodologic quality, the overall percent agreement was 88.5% (Kappa 0.84, 95% CI, 0.72-0.96). We added this information on Page 8.

7) Critical appraisal: No risk of bias assessment, while this is encouraged by the Cochrane collaboration (usual standard for effectiveness systematic reviews). Authors stated "The EPHPP tool can be used to evaluate multiple study designs, including RCTs, case controls, cohort and cross-sectional studies."; however, some of these designs can hardly help for an effectiveness review. Thus, this sentence is puzzling. In addition, authors stated "For the purpose of this review, the authors a priori decided to only report on studies with a strong or moderate global rating." This is discouraged by the Cochrane collaboration. All relevant studies are usually included and results synthesized; specifically, results of the appraisal are of great interest as they can lead to rich sensitivity analysis (comparing lower vs. higher quality studies for instance).

Response: The EPHPP tool includes items related to risk of bias and includes items related to selection bias, study design, confounders, blinding, data collection methods, and withdrawals and dropouts. We have included this information on Page 8 of the manuscript. All 39 relevant studies are included in the reference list. We decided that given the heterogeneity of the 26 weak studies, in terms of the study participants and outcomes, we would report on the 8 moderately to strongly rated studies that evaluated clinical outcomes that could be attributed to the toolkit.

We also add in the Discussion on Pages 18-19: One of the limitations of the study was that we focused on studies that evaluated the effectiveness of a toolkit to support the integration of evidence into clinical care; therefore, the studies included in this review reported quantitative results.

8) Data extraction: The following sentence is very confusing: "Because many studies embedded the toolkit into a multi-faceted intervention and did not evaluate the toolkit separately, information regarding all of the components of the study intervention was extracted". Thus, what was assessed? This seems to confirm the lack of conceptualisation.

Response: We have revised this sentence on Page 9:

Because many studies embedded the toolkit into a multi-strategy intervention (i.e., toolkit with an additional intervention) and did not evaluate the toolkit separately, information regarding all of the components of the study intervention was extracted. As well, the type of evidence, if any, underpinning the development of the toolkits' clinical contents and choice of resources was extracted.

9) Synthesis: The data analysis must be specified. There is no tabulation of quantitative results, no forest plot. Authors did not mention that a meta-analysis was not feasible. Results seem to report some type of narrative-like synthesis, while there is no clear narrative synthesis method.

Response: We have added a section in the Methods to address this point on Page 9:

When data from studies were available (e.g. means, standard deviations, proportions), meta-analyses would be conducted. A weighted mean difference (WMD), or a standardized mean difference (SMD), relative risk (RR), risk difference (RD) all with 95% confidence intervals (CI) would be conducted using a fixed effects model. If pooling of results would not be possible, a narrative descriptive review of study results would be presented.

On Page 10 of the Results section, we state the following:

Given the diversity of studies in terms of the study participants and outcomes, a meta-analysis was not possible; therefore, we chose to report on all studies with a strong or moderate global ratings rather than focusing only on RCTs of potentially weak quality.

VERSION 2 – REVIEW

REVIEWER	Alison Kitson University of Adelaide South Australia Australia
REVIEW RETURNED	18-Feb-2015

GENERAL COMMENTS	<p>The authors have done well to navigate themselves through a difficult and conceptually confusing landscape. In order for the paper to help guide future research activities, i would suggest the authors consider strengthening their argument in the following areas:</p> <ol style="list-style-type: none"> 1. what is the relationship (if any) between a 'multifaceted intervention' and a 'toolkit'? 2. who determines what goes into a 'toolkit'? 3. what are the challenges of evaluating the effectiveness of a 'toolkit' if it is embedded in a larger multi-strategy intervention? 4. Is there any way from your review that you could classify the main types of elements that make up a 'toolkit'? 5. In the absence of a working definition of 'toolkit' I would suggest you promote your definition and have this as one of your conclusions. However, you may want to re look at this and compare this definition to what you found in your results as i think the definition may be broadened 6. Another 'gap' from my perspective is around who 'facilitates' the 'toolkit' contents. i am not clear from the paper whether this was missing from the original studies or that the team did not consider it important. 7. Your paper reflects the current lack of real understanding around the development and implementation of KT interventions and again I think your conclusions should be promoting a more systematic approach to creating more clarity. The four points you have articulated could be strengthen around the types of common elements (education, feedback etc) and delivery methods
-------------------------	---

	<p>(facilitation, self-directed learning, reminders) and evaluation approaches (combining outcome and process measures including context)</p> <p>Otherwise, an important contribution to this challenging area of research.</p>
--	---

VERSION 2 – AUTHOR RESPONSE

1. What is the relationship (if any) between a 'multifaceted intervention' and a 'toolkit'?

Response: A variation on multifaceted KT interventions is the toolkit. Toolkits offer greater flexibility of use, and for the purposes of this review, are defined as a packaged grouping of multiple KT tools and strategies that codify explicit knowledge (e.g., templates, pocket cards guidelines, algorithms), and are used to educate and/or facilitate behaviour change.[9] Use of KT strategies housed within a toolkit are not necessarily prescribed in any combination or temporality (e.g., Strategy A +/- Strategy B +/- Strategy C, etc). The goal is for the user to select KT strategies in the toolkit that are supported by evidence of effectiveness and for use at their own discretion, according to their aims, resources, and context. Toolkits differ from multifaceted interventions in which the coupling of more than one KT strategy must be implemented together to comprise the 'KT intervention'; e.g., Strategy A + Strategy B = multifaceted KT strategy.

2. Who determines what goes into a 'toolkit'?

Response: The toolkit developers would determine the toolkit content. Again, the goal would be to include KT strategies that are supported by evidence of effectiveness and that are tailored to their aims, resources, and context.

3. What are the challenges of evaluating the effectiveness of a 'toolkit' if it is embedded in a larger multi-strategy intervention?

Response: One of the challenges of evaluating the effectiveness of a toolkit if the toolkit is embedded in a larger multi-strategy intervention, is the degree to which the study outcomes are influenced by the toolkit. The goal of any evaluation depends in part on (1) whether the toolkit elements themselves have document evidence of their effectiveness in isolation in previous studies; (2) whether one is evaluating the toolkit as a whole, and (3) whether the intervention is really a multi-strategy intervention (e.g., the authors are prescribing the toolkit + strategy "x" as their KT intervention).

4. Is there any way from your review that you could classify the main types of elements that make up a 'toolkit'?

Response: We report in our abstract and manuscript (Pages 12-13) that the types of resources embedded within toolkits varied but included predominantly educational materials. This is fairly eclectic and non-prescriptive, other than to hope that toolkits are comprised of evidence-based tools/strategies, but this is not always documented.

5. In the absence of a working definition of 'toolkit' I would suggest you promote your definition and have this as one of your conclusions. However, you may want to re look at this and compare this definition to what you found in your results as I think the definition may be broadened.

Response: We based our definition of a toolkit based on the scoping review conducted by Barac et al. (2014) on Pages 4- 5 of the manuscript:

"For the purposes of this review, toolkit was defined as packages of multiple resources that codify explicit knowledge (e.g., templates, pocked cards, guidelines, algorithms, and are used to educate and/ or facilitate behaviour change". We feel that this definition aligns with our results (See Table 1 for descriptions of the toolkits included in the review). We have added some text to provide clarity

regarding the distinction between toolkit and multifaceted KT intervention, as explained above, on Page 5 of the revised manuscript.

6. Another 'gap' from my perspective is around who 'facilitates' the 'toolkit' contents. I am not clear from the paper whether this was missing from the original studies or that the team did not consider it important.

Response: In Table 1, we include whether the toolkit described facilitation strategies (e.g., training sessions, local champions). Not every element of a toolkit requires facilitation. For example, in the study by Shah et al. (2012), the toolkit was sent to family physicians.

7. Your paper reflects the current lack of real understanding around the development and implementation of KT interventions and again I think your conclusions should be promoting a more systematic approach to creating more clarity. The four points you have articulated could be strengthened around the types of common elements (education, feedback etc) and delivery methods (facilitation, self-directed learning, reminders) and evaluation approaches (combining outcome and process measures including context).

Response: We appreciate this feedback and have incorporated this information in our conclusions on Page 17:

In summary, toolkits have potential as a promising KT strategy for facilitating practice change in healthcare. To fully understand their effectiveness, a systematic approach to planning and reporting their development, the evidence underlying each component, and any direction regarding appropriate implementation is required. Toolkits should have 1) a clearly described purpose, rationale for each component; 2) components that are rigorously developed and informed by high-quality evidence, such as systematic reviews; 3) delivery methods that are guided by a comprehensive implementation process (e.g., self-directed, facilitation, reminders) with consideration for fidelity of implementation where appropriate; and 4) a rigorous evaluation plan and study design that can help explain the factors underlying their effectiveness and successful implementation (i.e., combining outcome and process measures including context).[9]