Author's response to reviews

Title: Guidelines for the use of survivorship care plans: A systematic quality appraisal using the AGREE II instrument

Authors:

Sarah A Birken (sarah1@email.unc.edu)
Shellie D Ellis (sellis4@kumc.edu)
Jennifer S Walker (jennifer_walker@unc.edu)
Lisa D DiMartino (dmartin@email.unc.edu)
Devon K Check (dkcheck@live.unc.edu)
Adrian A Gerstel (gerstel@email.unc.edu)
Deborah K Mayer (dmayer@unc.edu)

Version: 3
Date: 25 March 2015

Author's response to reviews: see over
Dear Dr. Kent,

We thank you for this thorough, constructive review of our manuscript, "Guidelines for the use of survivorship care plans: A systematic quality appraisal using the AGREE II instrument." We appreciate your time, the reviewers’ time, and the opportunity to revise the manuscript.

We hope that our responses meet your expectations. If you find that our responses do not address your concerns, please let us know; we would be happy to attempt another revision.

Please find the text of your review below in bold. Our responses, as applicable, follow your review in normal font:

**Reviewer: Melissa Brouwers (Comments to the Author):**

The authors are interested in understanding why cancer survivorship care plans are not being used. As a first step, they investigated the quality of CSPs. If the quality is poor, improving their quality may be first step to improving their adoption. If the quality is good, then efforts could be directed towards more interventional efforts to increase their adoption.

1. The quality of the research questions and the rationale underpinning them are solid and well presented.

2. The methods used execute the study are appropriate, comprehensive, and well presented. Specifically, while concise, the search strategy, eligibility criteria, data extraction strategy (including supplementary file), and analytical framework descriptions were sufficiently detailed to enable the reader to follow precisely what occurred. The modifications made to the AGREE II items to better align with the research context were well justified and described. So too were the criteria used to categorize the guidelines as strongly recommend, recommend and do not recommend.

3. The paper would be strengthened by linking the results back to the implementation science literature and the thesis of the paper – why are CSPs not being used. For example, while most guidelines were not recommended, a few were though not strongly. Are they good enough?

   **Should efforts be directed to choose the best one or two candidates and optimizing those – from an implementation perspective, is that the best use of resources and effort?**

To more clearly link results back to the implementation science literature and the thesis of the paper, we have added the following paragraph to the discussion section:

“As we suggested might be the case in the introduction, the poor quality of guidelines for SCP use may contribute to limited and inconsistent SCP use in practice [3, 8]. Clear and unambiguous guidelines for SCP use may promote the effectiveness with which SCPs are implemented [6, 7]; in effect, high-quality guidelines are an implementation strategy [36]. In particular, guidelines that use behaviorally specific terms may be the most effective way of increasing implementation [6]. Guidelines for the use of SCPs may be improved with precise and consistent definitions of which templates are best to use; to whom the guideline applies; to whom the guideline does not apply; and when, where, and by whom SCPs may be most effectively developed and delivered. Clear specifications may facilitate evaluation of adherence
to guidelines. Future research should assess the relationship between guideline quality and the effectiveness of SCP implementation.” (p. 15)

The answer to your question about whether the few guidelines that we recommended were good enough is a definitive ‘no’ – at least not in isolation from each other. To leverage their collective qualities, we are currently drafting a manuscript which synthesizes information from the highest-scoring guideline for each AGREE II domain and creates a “meta-guideline” for SCP use. In the revised manuscript, we reference this effort as follows:

“Many of the guidelines included in this study scored highly in some, but not all, domains. Efforts to promote SCP use may benefit from combining elements from multiple guidelines into a single clear, unambiguous resource. SB, JW, LD, and colleagues are currently synthesizing information from the guidelines that scored highest in each AGREE II domain in this study to create a “meta-guideline” for SCP use that practitioners may reference to facilitate their decision regarding adopting SCPs and, should they adopt SCPs, to facilitate their implementation. By leveraging the work of existing guidelines, this effort represents an efficient approach to facilitating practitioners’ decisions around SCP adoption and implementation.” (p. 15)

Could an effort be integrated into a larger quality improvement agenda – can the implementation science speak to this more directly?

The intent of the paper is not simply to provide a summary of the quality of CSPs, but rather, how quality of CSPs may provide clues to why their adoption is not optimized. How does this study advance the research enterprise? How could these data be used to inform future questions?

In the revised manuscript, we make two suggestions about our study’s implications for broader quality improvement efforts. First, in the discussion section, we suggest that guidelines should be improved with behaviorally specific terms, and that the relationship between guideline quality and implementation effectiveness should be assessed in future research:

“As we suggested might be the case in the introduction, the poor quality of guidelines for SCP use may contribute to limited and inconsistent SCP use in practice [3, 8]. Clear and unambiguous guidelines for SCP use may promote the effectiveness with which SCPs are implemented [6, 7]; in effect, high-quality guidelines are an implementation strategy [36]. In particular, guidelines that use behaviorally specific terms may be the most effective way of increasing implementation [6]. Guidelines for the use of SCPs may be improved with precise and consistent definitions of which templates are best to use; to whom the guideline applies; to whom the guideline does not apply; and when, where, and by whom SCPs may be most effectively developed and delivered. Clear specifications may facilitate evaluation of adherence to guidelines. Future research should assess the relationship between guideline quality and the effectiveness of SCP implementation.” (p. 15)

In the conclusion, we suggest that efforts to improve guidelines should start in the guideline-development phase:

“Our findings indicate that the quality of guidelines for SCP use is poor. Guideline quality may be improved with more behaviorally specific information regarding methods for effective SCP implementation. Organizations that develop cancer-related guidelines may promote quality by
incorporating AGREE II into their formal practice guideline programs, as many other guideline-developing organizations have successfully done [15].” (p. 16)

4. The writing is excellent. Tables and figure are appropriate and informative.

Major Compulsory Revisions.

Link the results back the literature and the aim of the paper. See comment 4 above for a summary of the issues.

To more clearly link results back to the implementation science literature and the thesis of the paper, we have added the following paragraph to the discussion section:

“As we suggested might be the case in the introduction, the poor quality of guidelines for SCP use may contribute to limited and inconsistent SCP use in practice [3, 8]. Clear and unambiguous guidelines for SCP use may promote the effectiveness with which SCPs are implemented [6, 7]; in effect, high-quality guidelines are an implementation strategy [36]. In particular, guidelines that use behaviorally specific terms may be the most effective way of increasing implementation [6]. Guidelines for the use of SCPs may be improved with precise and consistent definitions of which templates are best to use; to whom the guideline applies; to whom the guideline does not apply; and when, where, and by whom SCPs may be most effectively developed and delivered. Clear specifications may facilitate evaluation of adherence to guidelines. Future research should assess the relationship between guideline quality and the effectiveness of SCP implementation.” (p. 15)

The answer to your question about whether the few guidelines that we recommended were good enough is a definitive ‘no’ – at least not in isolation from each other. To leverage their collective qualities, we are currently drafting a manuscript which synthesizes information from the highest-scoring guideline for each AGREE II domain and creates a “meta-guideline” for SCP use. In the revised manuscript, we reference this effort as follows:

“Many of the guidelines included in this study scored highly in some, but not all, domains. Efforts to promote SCP use may benefit from combining elements from multiple guidelines into a single clear, unambiguous resource. SB, JW, LD, and colleagues are currently synthesizing information from the guidelines that scored highest in each AGREE II domain in this study to create a “meta-guideline” for SCP use that practitioners may reference to facilitate their decision regarding adopting SCPs and, should they adopt SCPs, to facilitate their implementation. By leveraging the work of existing guidelines, this effort represents an efficient approach to facilitating practitioners’ decisions around SCP adoption and implementation.” (p. 15)

In the revised manuscript, we make two suggestions about our study’s implications for broader quality improvement efforts. First, in the discussion section, we suggest that guidelines should be improved with behaviorally specific terms, and that the relationship between guideline quality and implementation effectiveness should be assessed in future research:

“As we suggested might be the case in the introduction, the poor quality of guidelines for SCP use may contribute to limited and inconsistent SCP use in practice [3, 8]. Clear and unambiguous guidelines for SCP use may promote the effectiveness with which SCPs are implemented [6, 7]; in effect, high-quality guidelines are an implementation strategy [36]. In particular, guidelines that use behaviorally specific
terms may be the most effective way of increasing implementation [6]. Guidelines for the use of SCPs may be improved with precise and consistent definitions of which templates are best to use; to whom the guideline applies; to whom the guideline does not apply; and when, where, and by whom SCPs may be most effectively developed and delivered. Clear specifications may facilitate evaluation of adherence to guidelines. Future research should assess the relationship between guideline quality and the effectiveness of SCP implementation.“ (p. 15)

In the conclusion, we suggest that efforts to improve guidelines should start in the guideline-development phase:

“Our findings indicate that the quality of guidelines for SCP use is poor. Guideline quality may be improved with more behaviorally specific information regarding methods for effective SCP implementation. Organizations that develop cancer-related guidelines may promote quality by incorporating AGREE II into their formal practice guideline programs, as many other guideline-developing organizations have successfully done [15].” (p. 16)

Minor Essential Revisions.
Please add the database search strategies (perhaps an appendix), if available. It would be useful for anyone who may wish to update the search CSPs in the future.

Please see Figure 1 for the diagram describing our database search strategy and process. If, in addition to the reference to Figure 1, it would be helpful to elaborate on the search in the methods section, please let us know, and we would be happy to do so.

Reviewers: Joan Vlayen (Comments to the author):

Major Compulsory Revisions

1. I don’t agree with the removal or rewording of AGREE II items: - I do understand that SCPs are complex interventions, but also for these types of interventions, the link between evidence and recommendations should be explicit. Particularly in this case, the weak evidence base might in fact be the problem why SCPs are not frequently used.

- Furthermore, by changing items, the validity of AGREE II is changed, so this should really be avoided.

- Health benefits, side effects etc were considered not to be applicable to guidelines for SCP use, but still the authors use these kind of outcomes in the introduction section to support the use of SCPs.

We appreciate the reviewer’s concerns about the removal and modification of AGREE II items. We based our decision to remove and modify some AGREE II items on guidance provided in the AGREE II manual. Although it states that excluding items in the appraisal process is discouraged, the AGREE II manual offers the following guidance for when doing so is appropriate:

“On occasion, some AGREE II items may not be applicable to the particular guideline under review. For example, guidelines narrow in scope may not provide the full range of options for the management of the condition (see item 16). AGREE II does not include a “Not Applicable” response item in its scale. There are different strategies to manage this situation including having appraisers skip that item in the
assessment process or rating the item as 1 (absence of information) and providing context about the score. Regardless of strategy chosen, decisions should be made in advance, described in an explicit manner, and if items are skipped, appropriate modifications to calculating the domain scores should be implemented” (p. 12 of AGREE II manual).

We chose the recommended strategy of skipping the item in the assessment process. The reviewer’s concerns may be assuaged by the comments of Reviewer 1, principal investigator of the AGREE Next Steps Consortium and the AGREE A3. She said the following regarding our removal and modification of AGREE II items:

“The modifications made to the AGREE II items to better align with the research context were well justified and described.”

Please also note that, to some extent, item 11 “The health benefits, side effects, and risks have been considered in formulating the recommendations” and 16 “The different options for management of the condition or health issue are clearly presented” are captured in other parts of our analysis. For example, under the “clarity of presentation” domain the NCCN guideline states that, based on lack of evidence, SCPs are recommended but not required. We might infer from this that the alternative management option is not to use SCPs.

We acknowledge that lack of clear evidence of SCPs’ benefits may impede SCP use. However, this point may be beyond the scope of this paper: AGREE II does not suggest that guidelines should only recommend innovations for which evidence is clear; instead, AGREE II suggests that guidelines should (1) use systematic methods to search for evidence (p. 2); (2) clearly describe criteria for selecting evidence (p. 2); and (3) explicitly link recommendations to the supporting evidence (p. 3). Our revision of item 12 required that “evidence [be] described, and recommendations follow from it” (p. 9). In our revised manuscript, we acknowledge that a key weakness of guidelines for SCP use is their failure to link recommendations to evidence:

“Guidelines were generally definitive in their recommendations to use SCPs, but these recommendations were often not explicitly linked to evidence.” (p. 14)

To further address the reviewer’s concerns, we replaced references to ‘evidence-based practices’ with ‘recommended practices’ whenever possible (e.g., “Guidelines are tools that are intended to promote the use of recommended practices” (p. 5)).

2. The authors state in their conclusion: ‘Our finding that SCP use guideline quality is low is consistent with that of Reames et al’. I believe this is an inconsistency in this article. The authors believe that not all AGREE II items are applicable to the SCP use guidelines, but do compare with clinical cancer guidelines where the full applicability of AGREE II is out of question. Furthermore, these conclusions are not really conclusions, but still part of the discussion, because new elements appear.

Our primary purpose in including this point was to emphasize that our findings are consistent with other studies that have reported poor quality of cancer care guidelines. For clarity, we have revised this statement to read, “Overall, AGREE II domain scores suggest that the quality of guidelines for SCP use is low. This finding is consistent with that of other studies that have shown poor quality of guidelines in cancer care [34]” and moved it to the first paragraph of the discussion section (p. 14).
To address the reviewer’s concern about the conclusions section presenting new information, we have revised it as follows:

“Our findings indicate that the quality of guidelines for SCP use is poor. Guideline quality may be improved with more behaviorally specific information regarding methods for effective SCP implementation. Organizations that develop cancer-related guidelines may promote quality by incorporating AGREE II into their formal practice guideline programs, as many other guideline-developing organizations have successfully done [15].” (p. 16)

3. In general, the authors seem already to be convinced about the effectiveness of SCPs. However, the evidence that I saw is not convincing at all, and in fact mainly focused on breast cancer. If the authors are really believers, they should give stronger arguments in the introduction and/or discussion on why care providers should use SCPs. Again, the weak evidence base might be the reason for ‘underuse’.

We are interested in SCP use because, as we describe in the introduction, despite widespread recommendations for SCP use, implementation is limited. Our beliefs about the effectiveness of SCPs are beyond the scope of this paper. In reviewing the introduction, however, we can understand how the reviewer might have interpreted our review of the evidence as advocacy for SCP use. To address this, we have removed the statements in the introduction regarding effectiveness (i.e., “Three randomized controlled trials found no influence of SCPs on key survivor outcomes [3-5]; however, several observational and quasi-experimental studies demonstrated improved communication between survivors and among their providers, quality of follow-up care, and health outcomes when survivors receive SCPs”). Also, as stated in response to the reviewer’s second comment, we now acknowledge in the discussion section that a key weakness of guidelines for SCP use is their failure to link recommendations to evidence: “Guidelines were generally definitive in their recommendations to use SCPs, but these recommendations were often not explicitly linked to evidence.” (p. 14)

4. What I missed in the analysis is an overview of the actual recommendations, and how they relate to the AGREE II assessment. I cannot imagine that all identified guidelines recommended the same. Assessment with AGREE II does not tell you anything about the correctness of these recommendations, and I think the reader should be warned for that.

We believe that the reviewer is suggesting that we summarize the guidelines’ recommendations (i.e., to use SCPs or not to use SCPs) and offer a disclaimer that our designation of a guideline as “recommended” does not necessarily suggest that we advocate the guidelines’ recommendations. If we understood the reviewer’s comment correctly, we believe that doing so is beyond the scope of this paper. The AGREE II manual indicates that “the AGREE instrument is a tool that assesses the methodological rigour and transparency in which a guideline is developed” (p. 1). AGREE II evaluation does not offer insight into the quality of a guideline’s recommendations. In our view, including a disclaimer indirectly suggests that AGREE II in some sense offers insight into the quality of guidelines’ recommendations when, in actuality, it does not.

If we have misunderstood the reviewer’s comment, or the reviewer has a suggestion for addressing her concern without suggesting that AGREE II in some sense evaluates the quality of guidelines’ recommendations, we would be happy to discuss this issue further.

Minor Essential Revisions
5. In the introduction the authors say 'key survivor outcomes', I would say 'key patient-reported outcomes'.

This term has been omitted in response to the reviewer's first set of comments.

6. How many reviewers scored the guidelines? Was it two (DC and LD) or three (also SB)?

In the original manuscript, we used the term "score" to refer to steps taken by both DC and LD and SB. We understand how this may have contributed to the reviewer's confusion. To address this concern, we have replaced the term “score” in reference to steps taken by DC and LD to evaluate the guidelines with the term “rate” as follows:

“Each item is rated on a seven-point Likert scale, with 1 assigned for items with no clear discussion and 7 for exceptional quality of reporting. DC and LD read the entire AGREE II user’s manual and then independently rated all included guidelines” (p. 8).

SB scored the guidelines: “SB conducted final scoring, according to the instrument protocol, by adding together DC and LD’s respective ratings for items in each domain and standardizing the total score out of 100%” (p. 9).

We apologize for using the term “score” for both of these analytical steps.

7. Two out five recommended guidelines fail to have good scores on domain 3. I don't consider these guidelines to be of sufficient quality. Purely focusing on the 50% rules seems dangerous to me, and that's in fact why AGREE II does not provide rules about interpretation.

Despite lack of AGREE II rules for interpretation, we made the decision to offer categorical recommendations to promote consistency with extant studies using the AGREE II instrument and to clearly and unambiguously identify potentially useful guidelines for SCP use. In so doing, we hoped to facilitate practitioners’ evaluation of the quality of the many guidelines for SCP use. We appreciate the risk in this kind of categorical evaluation because, for example, practitioners may not have the time to understand nuances of our analytical method. As such, we were even more conservative than the authors whose methods we leveraged. Guo et al. (2014), for example, “strongly recommended” guidelines if the majority of domains (more than 4) scored above 50%; we required all domains to score above 50% to be strongly recommended. Huang et al. (2013) dichotomized recommendations, deeming any guideline “satisfactory” if it scored at least 50% in all domains; by grading our recommendations, we provide additional nuance.

The reviewer’s concerns may also be assuaged by the commendation of the Reviewer 1, principal investigator of the AGREE Next Steps Consortium and the AGREE A3: “[T]he criteria used to categorize the guidelines as strongly recommend, recommend and do not recommend...[were well justified and described]” (comment 2).

Discretionary revisions:

8. The G-I-N library was not searched: why?

We apologize for this oversight. Upon reading this comment, we searched the G-I-N library to address the reviewer’s concern and identify any guidelines that we might have included had we initially searched
it. We found only two results. We would have excluded the first (Models of care for cancer survivorship. Program in Evidence-Based Care NGC: 009571. Agency for Healthcare Quality and Research) because it does not include recommendations regarding the development, delivery, and/or use of SCPs during follow-up care. We included the second (Cancer Survivorship Care; Comprehensive Cancer Center of the Netherlands) in our analysis because we identified it in another electronic database.

We look forward to your final decision.

Sincerely,

Sarah A. Birken
Shellie D. Ellis
Jennifer S. Walker
Lisa D. DiMartino
Devon K. Check
Adrian A. Gerstel
Deborah K. Mayer