Article details: 2019-0193	
Title	Assessing quality of clinicians' conversations with patients and families before and after implementation of the Serious Illness Care Program in a hospital setting: a retrospective chart review
Authors	Christina Ma MD, Lauren Riehm MD, Rachelle Bernacki MD MS, Joanna Paladino MD, John J. You MD MSc
Reviewer 1	Mr. Bryan B Franco
Institution	School of Medicine, Queen's University, Kingston, Ont.
General comments (author response in bold)	1) This is a retrospective chart review performed to assess the effectiveness of a quality improvement initiative regarding improving serious illness conversations in the inpatient setting. This is a timely paper to address a clear research question using appropriate methods, however, it can benefit from clarification of methods and better acknowledgement of major intrinsic limitations. **We have added more text regarding the methods and limitations of the study in the revised paper (including the points indicated in the above and below responses to editor and reviewers' comments) (Pages 4-5, 9).** 2) a. The methods were sufficiently described but could be clearer by including additional information. First, it was unclear how the control group was created from the 2 cross-sectional sampling exercises. Was there a random process to establish the control groups? Was there an attempt to match participants? Similarly, how was the convenience sample determined (e.g. by the Unit Champion? The clinicians?)? **In 2 cross-sectional sampling exercises, we identified in real time all prevalent inpatients who had been hospitalized for at least 48 hours on the Internal Medicine service at Hamilton General Hospital on a single arbitrarily chosen day in December 2016 and on a single arbitrarily chosen day in February 2017. We then included in our control group those patients who met the SICP eligibility criteria (i.e., interRAI ED screener scores of 5 or 6). This is described in the Methods section (Page 5). The 2 cross-sectional sampling dates were chosen by convenience, not at random, and our study was not a randomized comparison of intervention and control patients. We did not attempt to match participants by other characteristics.** **Regarding the description of our convenience sample, please see our response to Editor comment #7 above.** b. Second, the results include data on which type of clinician (attending vs resident/nurse practitioner) led the conversation. The methods by which this data was collected is not described—was it based on who dictated the note in the chart? **These data were obtained based on who was explicitly stated in the note to have led the conversation. If this information was not included, we based it on which type of clinician signed the note. This has been added to the Methods section (Page 6).** c. Thirdly, were all attending, residents, and nurse practitioners on the unit trained in the SICP for the intervention? **Please see our response to Editor comment #5 about which clinicians were eligible for training and how many took part in the training. Overall, 21/26 (81%) of eligible clinicians were trained. This information has been added to the revised manuscript (Page 4).**

	d. Finally, in the control group, it is stated that "many charts had multiple documented conversations that scored one point or greater." Would it have been more of a fair comparison to score the compilation of these charts (since it would presumably have covered more points) rather than the highest scoring conversation? Do serious illness conversations have to occur at one time or is there a draw back to having more than one, as it appears to be the case in the control group? Clarifying these questions can add the methodological rigour of this paper. **Please see our response to editor comment #10.** 3) The project is described as a quality improvement (QI) initiative, but the manuscript lacks description of many aspects relevant to QI. For example, there is no discussion of process outcomes (e.g. a measure of clinicians' knowledge/attitudes after attending the SICP, were there any clinicians on the unit that did not receive training in SICP, and if so, what were the reasons?). There was also no mention of balancing measures—did the intervention have negative/positive/neutral effects on other functions of the inpatient unit? E.g. how did the efficiency of the unit change after implementation? Furthermore, contextual considerations appear to be lacking. E.g. was there a culture shift in the hospital that helped with the uptake of the SCIP? how/if stakeholders were engaged? were there lessons learned from implementation? A description of these QI-related factors can be valuable in understanding the SCIP as a quality improvement initiative and how it can be adapted in other settings. The SQUIRE 2.0 guidelines may be useful in guiding this discussion. **Please see our response to editor comment #4.** 4) A more transparent discussion of the study's limitations can strengthen this paper. First, given the unblinded and retrospective nature of the study, there are many confounders. For example, you note that the SCIP appears to have resulted in better documentation—the primary outcome (quality of conversations) was measured using charts and therefore, the differences between the two groups may be attributable to better documentation. What are other potentially important unmeasured/unreported differences between the two arms (e.g. different diagnoses, different clinicians, etc.)? Finally, there appears to be a lack of emphasis on patients' perspectives which is particularly important since the purpose of the project was to improve person-centred care. The primary outcome of "high quality conversations" was studied from researchers and clinicians' subjective perspectives. These limitations suggest important avenues for future research and likely warrant a discussion in this manuscript. **We agree and have added more text regarding limitations of the study in the revised paper (including the limitations indicated in the above points) (Page 9). We also agree that the differences in group may be attributable to documentation however, there is no way to rule this out as there is no way to know about conversations that occurred but were not documented (especially in the control group). Documentation is the only way to retrospectively identify conversations that took place. Again, we have included this in our limitations section. Finally, we agree this paper did not emphasize the patient perspective, however this is a targeted assessment of the quality of conversations based on retrospective chart review and a future paper is planned that will report on a range of outcomes of our multi-site implementation of SICP (see also responses to Editor comments #4 and #5),**

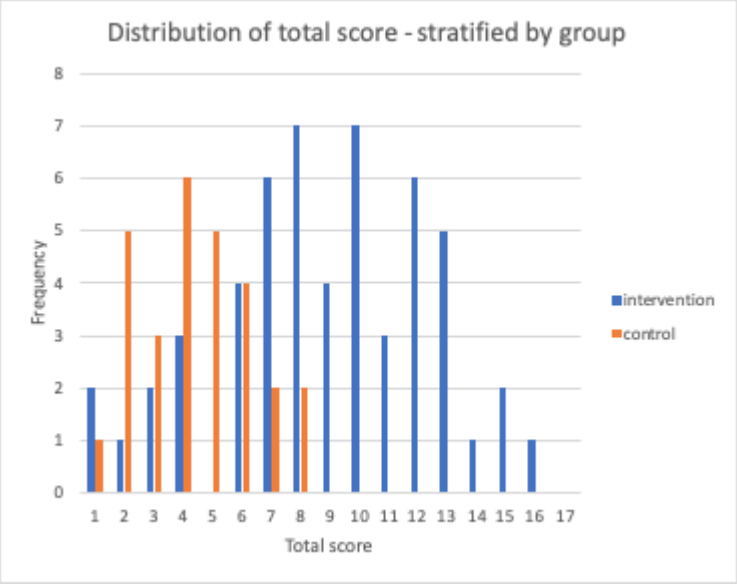
	including patient-reported outcomes and clinician-reported outcomes. 5) In the Discussion and Conclusion, the implications of the study are unclear. First, the SCIP was implemented with a newly hired Unit Champion. Is this sustainable? Are there plans to for a cost-effectiveness analysis for this intervention? Secondly, are there implications for clinicians moving forward; e.g. will all future clinicians in the inpatient unit and elsewhere in the hospital participate in SCIP? Finally, how does SCIP fit within the context of the health care system/hospital. There is a push for "value-based" health care under which this intervention can arguably be included. Literature also suggests that targeting clinician behaviour through education has a limited impact—systemic changes (e.g. to payment models) are needed to sustain and promote improvement. E.g. will or should SCIP affect payment models for clinicians working in the unit? These are just a few ideas of how the study's implications can be discussed by the authors and by no means is an exhaustive list. Including these in the discussion can help readers understand the study within the context of their own institutions and in the QI and person-centred care literature. **The implications of our study are that the SICP is transferable and adaptable to a hospital setting, and that it appears to add value, as it increases the quality of serious illness conversations and adherence to best practices compared to usual care. We agree with the reviewer that system changes are critical to effect change and that training alone is unlikely to be enough--this study supports this notion since the SICP intervention was a multi-component program that included not only training, but specific, evidence-based tools and a robust system change component. We also agree with the reviewer that sustainability and cost-effectiveness are important issues and have included this in our revised Discussion section as an avenue for future work (Page 10). Again, the aim of this study was not to report in detail on the implementation itself or to assess the sustainability of the program, but rather to specifically compare quality of conversations during implementation versus usual care. We do have plans to report on the aforementioned issues, including that of sustainability, in future manuscripts. A more robust description of the implications of our study that include the above have been added to our Discussion section (Page 10).**
Reviewer 2	Dr. Braden Manns
Institution	Departments of Medicine and Community Health Sciences, Foothills Medical Center, University of Calgary, Calgary, Alta.
General comments (author response in bold)	This is an interesting article documenting a local experience in one hospital which implemented the Serious Illness Care Program (SICP) on the medical wards of our hospital to build capacity to have more frequent and higher quality serious illness conversations. As I reflect on the questions on this guide, it reminds me how important these conversations are – and that although I think I do a reasonable job of these discussions, I definitely don't go into the level of detail recommended by the SICP. So this initiative seems very important. The strength of this study is the initiative itself – using a patient screener to identify patients at risk of longer hospital stays and need for community resources; rolling out training sessions with physicians around the SICP; ensuring a structured conversation happened (and having it recorded in a structured way) and having some dedicated resources to get this started (a nurse champion). As currently reported, the weaker part of this paper related to the evaluation. I have some specific feedback that is hopefully helpful to the authors:

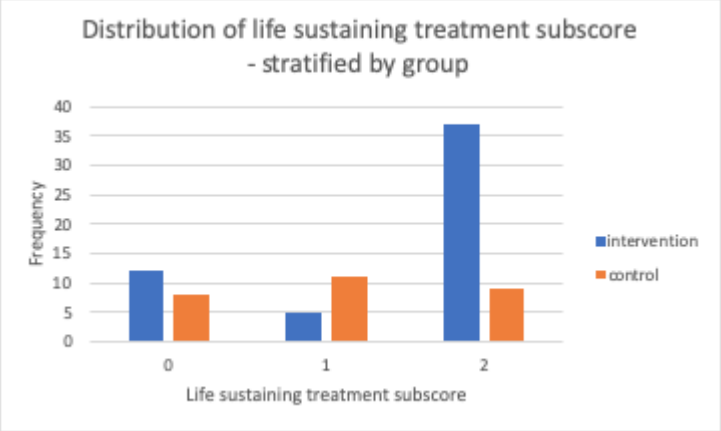
	1. Additional information would be helpful on this local initiative, of relevance to other hospitals considering such an initiative. How many physicians were trained (out of how many – eg how easy was this to roll out?)? How long does an "average" serious illness conversation last – and was this measured in this study? It seems like this might have been quite an expensive intervention, as a nurse was used to prep patients; and also to book conversations, etc. Is the intervention sustainable without this type of resource? Has this program been sustained? Is there any data on this? **Please see our response to Editor comment #5 regarding the number of clinicians trained (and the total eligible clinicians).** **Unfortunately, we did not time the serious illness conversations so cannot report on this variable. We have some rough estimates in addition to anecdotal experience suggesting that the average conversation probably takes 30 minutes but do not feel that this information is of sufficiently high enough credibility to include in the manuscript itself unless the editors request it.** **This paper is a targeted assessment using a chart review to compare the quality of conversations during the SICP implementation to that of historic controls at Hamilton General Hospital. As such, The purpose of this paper is not to report the entire implementation experience of the SICP or other related outcomes (such as sustainability). We do have plans to report a more detailed description of the (multi-site) implementation and comment on the sustainability of this intervention in future manuscripts (see also our response to Editor comments #4 and #5).** 2. You note that you modified the serious illness conversation / program, and the reasons you listed for the needed changes were reasonable (modified to the hospital setting and for the local setting), but I wasn't clear what changes were actually made, and whether any substantive changes made? Please clarify. **Please also see our response to editor comments #5 and #12d. The main adaptations for the hospital setting were: (i) to leverage an existing tool being used at our hospital (interRAI ED screener) to identify suitable patients, rather than asking clinicians on the ward to use the "surprise question" (would you be surprised if this patient died in the next year?) as in the original program; (ii) for the Unit Champion to prepare patients/families in person on the medical ward for a Serious Illness Conversation, rather than mail a pre-visit letter as was done in the original program; (iii) for the Unit Champion to remind clinicians (in person or by phone call/text message) of an upcoming Serious Illness Conversation, rather than email as was done in the original program; (iv) documentation of conversations was done by dictating a structured summary, since our hospital did not have an EMR for electronic documentation of clinical notes, as opposed to the EMR used for the original implementation at the Dana-Farber Cancer Centre which did have this functionality.** 3. Moreover, you note that you used a tool adapted from Paladino, Bernacki et al. (ref 21) to track the quality of the conversation that was documented (what is described as a validated codebook). This is described in Online appendix 2, but it is unclear what modifications were made? And after the modifications, is this still considered validated? **We did not make any modifications to the codebook and therefore this**

	codebook is still considered validated. We erroneously used the word "adapted" in the header for Online Appendix 2 and have changed this to read "From Paladino, Bernacki et al" 4. It's important to understand what changes you made to the SICP and the codebook, since significant changes may mean that neither instrument is valid. **Please see above responses to Reviewer 2 comments #2 and #3.** 5. Information within paragraph 2 of Online appendix should be available in the paper (specifically the overall screening initiative). **We agree and have moved this text from Online Appendix 1 to the main text in the methods section (Page 4-5). (See also our response to Editor comment #5 and #12d).** 6. More details in table 1 are required about patient baseline characteristics – what were the most common admitting diagnoses; what was the average length of stay. Aside from the charlson comorbidity information, do you have any indication of how acutely unwell these patients were. This is important as the end-of-life conversation, which did not commonly occur, may not have been relevant in many patients. **We agree and we have added more baseline characteristics to Table 2 (previously Table 1) (Page 7). These characteristics include the number of emergency department visits in the previous year, arrival in the emergency department by ambulance, being seen in cancer clinic in the previous 6 months, and various lab values.** 7. Do you have any sense as to whether you are really capturing the quality of the discussion by reviewing the chart notes only (vs simply good quality charting). It would be interesting to see how this correlates with the actual quality of the conversation, and perceptions from patients. **The reviewer raises an interesting point. However, there is no way to know about or assess the quality of conversations that occurred but that were not documented (especially in the control group). Having said that, it is certainly our anecdotal experience that it is primarily the quality of conversations, rather than only the quality of documentation, that improved after SICP implementation; however, we do not feel comfortable reporting our anecdotal experience in the published paper (unless the Editor requests it). Documentation is the only way to retrospectively identify and assess conversations that took place. We have included this in our limitations section (Page 9). Finally, we agree with the reviewer that patient perspectives are important. However this paper is a targeted assessment of the quality of conversations based on retrospective chart review and a future paper reporting on our subsequent multi-site implementation of SICP is planned that will include patient experience/patient-reported outcomes.** 8. The conversation only happened with 20% of patients who were eligible to have a conversation. One might argue that presenting your data the way you do is like analyzing an RCT using strictly a "per protocol analysis" – rather than the "intention to treat" analysis that is considered less biased. (You mention that this was because a clinician was limited in having too many of these conversations in any given week, which raises questions as to whether it's a practical tool for in-hospital use?). I'm curious whether this type of conversation occurred in other

	patients during the intervention time, but because it wasn't been "observed", might have occurred with lower quality. **See also our response to Editor comment #9. We agree with the reviewer that our design and analysis is more analogous to a "per protocol" analysis than an "intention to treat' analysis. To be more like an ITT analysis, we could have included all patients from the control group (of which half had no documented conversation and would have scored zero on the codebook), and all the eligible patients from the intervention group who did not receive a formal Serious Illness Conversation (and would have been akin to "usual care" in the control group). Using this approach, mean codebook scores in each group would approximate some sort of integral of quality and frequency of conversations. We do not have the resources to go back and abstract data on the eligible patients who did not have a formal Serious Illness Conversation. We have pointed out in the limitations section of the Discussion that our analysis does not capture the frequency of conversations in each group and that our findings are specifically conditional on a conversation having occurred in each group (Page 9). In this way, to use the RCT analogy, our analysis is more like an assessment of "efficacy" under ideal conditions than "effectiveness" under real world conditions.** 9. Table 2 and Figure 1 present the same info – suggest you choose one method (I like the figure). One question that arises when reading Table 2 is whether the changes – while statistically significant – are clinically significant. **We agree and have moved Table 2 into the Appendix section (now Online Appendix 3). Unfortunately there is no data available that correlates an absolute difference in scores to other outcomes or clinical significance and we have pointed this out in the Interpretation section of our revised paper (Page 8).**
Reviewer 3	Dr. Shannon Marie Ruzycki
Institution	Department of Medicine, Foothills Medical Center, University of Calgary, Calgary, Alta.
General comments (author response in bold)	Summary: This paper reports on implementation outcomes for a quality improvement intervention intended to improve the rates, quality of discussion, and documentation of serious illness conversations for medical inpatients. The authors found that implementation of their serious illness conversation program successfully improved all of their metrics for this project. The authors provide justification for the importance of this work. Authors: 1. The authors state that this is a retrospective chart audit study (in the Methods and Discussion), but this study may be more accurately described as a pre/post non-randomized intervention design or as evaluation of an implementation. The word "retrospective" is redundant as all chart audits are retrospective. **Please see our response to Editor Comment #4. We have consistently described our study as a retrospective chart review and highlighted the before-after nature of the study design in our revised manuscript since we believe this is the most concrete and accurate description of the primary focus of this paper. However, if the editors prefer, we are happy to remove**

	the "retrospective" from the description of our study design. 2. This could be a pre-/post-non-randomized intervention design using historical controls. If so, the authors should note that an important limitation is the tendency of historical control studies to overestimate treatment effect, since most metrics tend to improve over time even without specific intervention. In this case, the outcome is the quality of documentation of the conversations (as an assumed surrogate outcome for higher quality conversations). **We agree and have added the above to our limitations section (Page 9)** 3. Alternately, these results may be presented as an implementation evaluation. Since the authors are reporting the uptake and use of the intervention rather than outcomes based on the intervention, I would suggest restructuring the Methods and Results with the aim of communicating the results as an evaluation of implementation of a quality improvement intervention. This may be most appropriate because the outcomes reported are about fidelity (implementation as intended) and feasibility (ability to implement an intervention) of the implementation. The authors could also comment on the appropriateness and adaptiveness of the intervention (e.g., how the intervention was adapted to the inpatient setting even though it was originally designed for the outpatient setting, and if additional changes needed to be made throughout in order to continue to adapt). If the authors decide to restructure, the StaRI reporting guidelines may be helpful. **Please see our responses to Editor Comments #4 and #5. We agree that these factors are also of interest to readers and plan to report this in a future manuscript describing our subsequent multi-site implementation of the SICP.** Major: Methods: 4. It would be helpful for readers, especially those unfamiliar with the SICP, if you summarized the components of the SICP in a table or figure. The "conventional" or usual components could be one column and your adaptions could be the third column. Since this paper describes implementation of the SICP, the actual protocol should not be in an Appendix, it should be described in the Methods, and should include a comparison to usual care. For example, this may help explain why the authors are reporting on where these conversations are documented (written versus electronic notes), and who had participated (attending physicians versus other). This would help the with the Discussion section, where you refer to specific domains of the codebook. **We agree and have moved the content of Online Appendix 1 to the main text in the Methods section of our revised paper (Pages 4-5). We have also added Table 1 as suggested by the reviewer to summarize the components of the SICP and the adaptations we made for the hospital setting (Page 4).** 5. The authors should confirm that the data analyzed is parametric, since they report means and standard deviations. Often scales with small amounts of numbers are non-parametric. **The reviewer raises a good point. Although the distribution of data for the total scores (maximum score of 17) is parametric (i.e. approximates a normal distribution), as shown in Figure A below, the distributions for some of the domain scores (e.g. maximum score of 2) were not parametric, as shown in**

	Figure B below. Even when data are non-normally distributed, the central limit theorem shows that sample means will approximate a normal distribution if sample sizes are sufficiently large (greater than 20 to 30) and that, as a result, parametric tests such as t-tests are still valid. **However, to take the most conservative and consistent approach for our revised manuscript, we have reported the median and interquartile range of codebook scores instead of mean and standard deviation, and have used a non-parametric test (Mann-Whitney) instead of a t-test to compare total scores and domain scores between the intervention and control groups. This has been updated in the Methods section of the revised manuscript and in the relevant main text, Figures, and tables in the Results section (Pages 7-8).** **Also, in the Online Appendix containing the scores for the individual codebook items in each group, instead of reporting these data as mean (SD), we have revised this to simply report the n (%) of patient charts in which the item was present or absent. We used a chi-squared test (or Fisher's exact test when cell sizes were small) instead of a t-test to compare these proportions (Page 8).** **Note that the overall results and conclusion remain the same despite these changes to the analytic approach.** **Figure A**  **Figure B**



6. Does the educational component for physicians focus on documentation? Are the aspects of documentation that are important in the codebook taught to the care providers?
Yes, the educational component includes brief instructions for clinicians to document serious illness conversations as dictated clinical notes that follow a structured format that mirrors the items in the Serious Illness Conversation Guide. Moreover, the focus of the educational component is to train clinicians to use the Serious Illness Conversation Guide; the elements/domains of the codebook mirror the items in the Serious Illness Conversation Guide.

7. Are the domains of the codebook related to the educational component of the SICP? Are the providers instructed to include all of the codebook components for documentation, or is the codebook unrelated to the SICP program?
Please see our response to Reviewer #3, Comment #6 immediately above.

Discussion:
8. The authors should note the limitations of comparing the intervention group to a historical control (e.g., overestimation of effect size). The authors should also note that higher quality documentation is being used as a surrogate of quality of conversations.
We agree and have included these limitations in our limitations section (Page 9).

Minor:
Methods
9. The interRAI Emergency Department Screener should be explained as it is in an inclusion criteria for the study and may not be familiar to all readers. This would include how it is used in your centre, which would be helpful for readers to understand if selection bias is possible. For example, do all patients receive this score at presentation?
We agree with the reviewer that this is important information. Originally, this was reported in Online Appendix 1 and we have moved this content to the main text in the methods section of our revised paper (Pages 4-5).

10. The language of the methods (page 7, lines 11-14) suggests that the study team members performed the serious illness conversation ("We screened 319 patients for eligibility, identified 275 eligible patients, and delivered serious illness

	conversations"). Is this true, or was it the attending medical team who was responsible for the conversations? **It was the attending medical team (attending physicians and nurse practitioners) who delivered Serious Illness Conversations. We have clarified this point in the revised paper (Page 5).** 11. It is not clear what the authors mean by the term "prospective" on page 7, line 20 of the Methods section ("we had prospectively conducted 2 cross-sectional sampling exercises"). Cross-sectional refers to one point in time - do the authors mean pre-planned? **Yes, we used the term "prospective" meaning "pre-planned". We have altered the text in the manuscript to clarify this (Page 5).** 12. Were the assessors masked to whether the patients whose charts they reviewed were in the intervention or control group? **Unfortunately, given the nature of the chart review and how the notes were documented, the assessors were not blinded to whether patients were in the intervention or control group. We have stated this in our limitations section (Page 9).** 13. The codebook could also be a figure or table in the main text, if spaced allowed. Domains could be summarized with an example item from each domain to give the reader a sense of how quality was assessed. **Unfortunately, space does not allow for the codebook to be included in the main text. We will already be appreciably increasing the length of the manuscript to address key concerns raised by the Editor and Reviewers herein so have judged that the codebook should remain as an online appendix. If the Editor feels strongly about this point and is willing to further increase the word limit of the manuscript accordingly, we are open to re-considering. In Online Appendix 2 containing the codebook, we have added in our revised submission an extra column that includes examples for each codebook item (Page 6).** 14. Similar to the interRAI, the authors should orient the scale used for the codebook (e.g., higher scores indicate greater compliance to high quality documentation). **We agree and have included this information in the methods section (Page 4-5).** Results: 15. The authors report whether the SICP note was dictated electronically or written for the intervention and control groups - is this a component of the SICP, and if so, why? If not, I would not highlight this information. **During the training sessions, clinicians were instructed to document serious illness conversations in the medical record by dictating a structured summary mirroring the items in the Serious Illness Conversation Guide. Therefore, we reported the number and proportion of serious illness conversations that were dictated into the electronic chart since it was a component of the SICP and because we believe dictated notes are more legible and more retrievable than handwritten notes.** 16. Did any patients decline the offer of a SICP?

	Please see our response to Editor comment #14. Discussion: 17. The authors mention the result that "all of these transcribed notes contained additional information that was not directly pertinent to the serious illness conversation" for the first time. No new results should be presented in the Discussion section. This should be mentioned in the Results if the authors wish to discuss it later. **We agree with this and have included this result in the Results section (Page 7).** Appendix/Figures: 18. The authors could include a third column for "Examples" for Online Appendix 2 instead of including the examples within the definition columns. **We agree and have added a third column for Examples in Online Appendix 2.** 19. Instead of "/7" the authors could consider "n=7" items. **We thank the reviewer for the suggestion. We think that using the "n=7" notation may potentially confuse readers since this type of notation often represents the number of subjects, which it does not in this case.** 20. The authors don't need to report p-values for differences in Table 1, as the differences are expected to be random. **We agree that this would be true for a randomized design, but since the subjects in our study were not allocated at random to the intervention and control groups, the P-value is actually pertinent in this case. Therefore, we have not removed the p-values.**
Reviewer 4	Mrs. Emily Mulligan
Institution	Winchester District Memorial Hospital, Research
General comments (author response in bold)	A very relevant and important topic as we talk more about patient-centered care. No doubts about originality. 1. Introduction: What triggered this QI project? What was your burning platform? Would like a better sense of the 'before' the intervention. **We believe that we adequately presented the rationale for implementing the SICP in the Introduction. We did make some minor edits to the Introduction section in our revised paper to further emphasize the rationale for our project (Page 3). If the Editor feels we should further revise this section, we are open to considering this.** 2. Setting: What are the patient volumes? **The average number of patients admitted monthly to the Internal Medicine program at Hamilton General Hospital in 2017 was 388. (Note that the number of eligible patients for SICP would be appreciably smaller since this number would include many patients who are admitted for less than 48 hours or who were younger patients without serious illness). We have added the information about patient volumes to the Setting section of the manuscript (Page 4).** 3. Intervention: How long did the intervention take to implement? A bit more info on timeframes.

	We have included information about the time frame for this study in the Methods section (see our response to Editor comment #3) (Page 4-5). 4. In the last paragraph of 'Evaluating quality of conversations' did you mean to put Appendix 2 instead of Appendix 1? **Yes, this was a mistake. However, since we are moving Online Appendix 1 to the main text, as suggested by the Editor and Reviewer comments, Appendix 2 is now Appendix 1 in the revised paper. We have ensured that all references to Appendices in our revised manuscript are now correct.** 5. First section of results, 'Serious Illness Conversations' could be shortened, it is a bit of a repetition of the table. **In our revised manuscript, we have tried to minimize repetition in the main text of the Results section of what is already reported in the Table(s).** 6. Is there a reference that you can cite when you say the potential outcome of improved documentation is to impact future care? **We have included the following reference in our Discussion section: Heyland DK, Ilan R, Jiang X, You JJ, Dodek P. The prevalence of medical error related to end-of-life communication in Canadian hospitals: results of a multicentre observational study. BMJ Qual Saf. 2016 Sep 1;25(9):671-9.** 7. What are the lessons learned that could be used by other sites? **Please see our response to Editor comment #19.** 8. Could you comment on the generalizability of the interventions? This might be hard for a smaller institution or somewhere rural that does not have the resources, financially or personnel, to for example, hire a unit champion. Also leveraging an existing initiative to screen patients might help with resources, but can also decrease generalizability. **The reviewer raises a good point. However, as mentioned above, we are planning a future manuscript that will directly address issues related to our experiences with the implementation of the SICP at multiple sites. Sustainability and generalizability are issues we plan to discuss in the future paper. In addition, we have mentioned in the "limitations" section of our revised paper that our findings may not be generalizable to other settings since ours was a single-centre study (Page 9).** 9. Would also include in the limitations the comment around the practical limitations of the clinicians to be able to hold more conversations. **We agree and have added this to the limitations section (Page 9).** 10. Would like more info on the implementation itself. Were there any other behavioural interventions used? What change management methods were used to deal with resistance to the change in culture and to ensure sustainability? **A brief description with more information on the implementation, which was previously found in Online Appendix 1, is now included in the Methods section (Pages 4-5). The purpose of this study is to report on a focused assessment of one outcome after the implementation of a QI project, not to report our entire multi-site implementation experience of the SICP or other related outcomes; this will be addressed in a future paper.**

	11. Concerning patient-centered care, was there any patient satisfaction observations, even if not formally measured? **This paper did not specifically assess the patient satisfaction outcomes. However, a future paper is planned based on our multi-site implementation that will include patient-reported outcomes.** 12. Were there any balancing measures? **Please see our response to Editor comment #4.** 13. Figure 1 did not print well, was hard to read **Thank you. We will ensure to include a higher resolution version of the Figure at the copy-editing stage if the paper is accepted for publication.**