PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Impact of Electronic Clinical Decision Support on Adherence to Guideline Recommended Treatment for Hyperlipidemia, Atrial Fibrillation, and Heart Failure: Protocol for a Cluster Randomized Trial
AUTHORS	Kessler, Maya; Carter, Rickey; Cook, David; Kor, Daryl; McKie, Paul; Pencille, Laurie; Scheitel, Marianne; Chaudhry, Rajeev

VERSION 1 – REVIEW

REVIEWER	Dean F. Sittig School of Biomedical Informatics University of Texas Health Science Center at Houston, TX USA
REVIEW RETURNED	18-Oct-2016

GENERAL COMMENTS	This manuscript reports on a protocol for a RCT trial of a new CDS intervention for 3 chronic conditions. I find the lack of references to similar studies of guideline adherence in the literature to be unacceptable. Many studies have been done over the last 20 years very similar to this one. Failure to recognize past attempts does not make your study innovative. Here are a few examples just to get you started. I think any new trial of an old technology should carefully review what has been tried, what worked, what didn't, and then there should be an explanation of why the current proposed trial has a chance to succeed. See for example: Tierney WM, Overhage JM, Takesue BY, Harris LE, Murray MD, Vargo DL, McDonald CJ. Computerizing guidelines to improve care and patient outcomes: the example of heart failure. J Am Med Inform Assoc. 1995 Sep-Oct;2(5):316-22. Murray MD, Harris LE, Overhage JM, Zhou XH, Eckert GJ, Smith FE, Buchanan NN, Wolinsky FD, McDonald CJ, Tierney WM. Failure of computerized treatment suggestions to improve health outcomes of outpatients with uncomplicated hypertension: results of a randomized controlled trial. Pharmacotherapy. 2004 Mar;24(3):324-37. Tierney WM, Overhage JM, Murray MD, Harris LE, Zhou XH, Eckert GJ, Smith FE, Nienaber N, McDonald CJ, Wolinsky FD. Can computer-generated evidence-based care suggestions enhance evidence-based management of asthma and chronic obstructive pulmonary disease? A randomized, controlled trial. Health Serv Res. 2005 Apr;40(2):477-97. Tierney WM, Overhage JM, Murray MD, Harris LE, Zhou XH, Eckert GJ, Smith FE, Nienaber N, McDonald CJ, Wolinsky FD. Effects of computerized guidelines for managing heart disease in primary care.

	J Gen Intern Med. 2003 Dec;18(12):967-76. Lester WT, Grant RW, Barnett GO, Chueh HC. Randomized controlled trial of an informatics-based intervention to increase statin prescription for secondary prevention of coronary disease. J Gen Intern Med. 2006 Jan;21(1):22-9. I think you should present the baseline adherence rate to these guidelines in this paper. I can't believe that you haven't already done that, so why not report it? Why didn't you stratify the practices based on baseline adherence to the guidelines? What if the intervention group has a higher baseline rate of adherence and therefore does not need the advisor? This could make your study fail, due to adverse patient/physician/practice selection. I think you should be measuring and reporting the percentage of patient visits in which the advisor was available and that the clinician actually looked at it. In my experience, systems like this have low utilization, especially after the initial excitement has worn off, which it most likely will after 3 months or so and then you will have 3 more months of data collection. I would still expect you to analyze your data with an "intention to treat" model, meaning that you must still include all the patients/clinics even if the physicians didn't use the system. I think you should also make several "outcome" measures on the patients to help ensure that the new guidelines don't inadvertently lead to increased patient visits, ED visits, hospitalizations, or deaths. While I believe the guidelines are good, if there is an error in any part of this system and the clinicians follow these guidelines you may have problems. I would hope to see some improvement in patient health as a result of following these guidelines, but with such a low expectation of improvement in guideline adherence (see below), this is certainly unrealistic. You should at least look for these changes. I find it hard to believe that a study of this complexity is worth doing for only a change in adherence from 2% to 4% (see page 8 line 32). While you may get statistical significance, that change does not seem clinically or practically significant at all. Unless, your baseline compliance is already over 75%. And if it is over 75%, then I would also question why one would put so much effort into a study for such a small benefit. Please try to explain, why you think an organization should expend so much time and energy for such a small improvement, when there are major problems with quality in our healthcare system. Page 9 line 8 HIPAA is spelled wrong!

REVIEWER	David Kaufman Arizona State University, United States I have no competing interests. In interest of full disclosure, I am a Mayo Clinic Research Affiliate. However, I do not collaborate or work with any of the authors.
REVIEW RETURNED	28-Oct-2016

GENERAL COMMENTS	This is a very well-written proposal that addresses an important topic regarding the efficacy of the use of decision support tool at the point of care for 3 important conditions in view to enhance adherence to the latest guidelines. The protocol is well designed and clearly articulated. The group has substantial expertise and experience in this area. They are also leveraging resources such AskMayoExpert knowledge base to facilitate their efforts. In general, they are well positioned to successfully execute this research. Most of my concerns have to with matters that need to be more fully addressed in the manuscript. The first problem has to do with the rationale. The authors state that “many interventions have been designed to address these gaps between evidence and practice, yet most achieve only minimal gains in guideline adherence”. What is it about those interventions that have yielded only minimal gains? What is wrong with the “received” or conventional approach to decision support or computer-based clinical guidelines? Why is there an expectation that the use of MayoExpertAdvisor would yield superior results? What differentiates it from the other tools/approaches? It would be worth developing this argument in a paragraph or so. Similarly, it would be useful to know more about the Level 3 and 4 validation efforts. What principles or findings guided the iterative redesign effort? It would also be useful to know a little more about the results of the pilot beyond general satisfaction of those who were surveyed. For example, what was deemed to be of value and what needed to be revised. How was the system designed to be integrated with clinical workflow? Why will physician assistants be included in the study and residents excluded? In a sentence, please explain how the system employs NLP to extract data that would be otherwise unavailable. Similarly, a couple of sentences on the design of the user interface would be informative. Is it possible to use log file data in lieu or in addition to the survey to derive efficiency measures in managing the 3 clinical conditions or use of risk calculators? This would have greater validity than self-report measures. Interviews with a small number of the study participants would be illuminating. Of course, the choice of methods is at the discretion of the authors. The guidelines workflow diagrams in appendices 2 and 3 are very hard to read. There is ample room to expand them on a page.

	The authors appropriately raise the limitation regarding generalizability given the use of proprietary systems only available to the Mayo Clinic. A follow-up sentence (or so) should articulate how this research can still make a substantive contribution to knowledge regarding clinical practice (or decision support). In other words, what will be the nature of the take home message to those who wish to achieve similar clinical goals related to improved adherence to guidelines but lack access to MayoExpertAdvisor.

REVIEWER	Susan McRoy University of Wisconsin-Milwaukee
REVIEW RETURNED	23-Jan-2017

GENERAL COMMENTS	This paper describes a protocol for evaluation. The motivation, methods and measures make sense and the presentation is clear and understandable. This reviewer is not a specialist is statistics so the appropriateness of the precise study size will need to be determined by another reviewer.

REVIEWER	Dympna O'Sullivan City, University of London, UK
REVIEW RETURNED	25-Jan-2017

GENERAL COMMENTS	Important work in an area where guideline adherence is poor. Comprehensive validation of the tool thus far and I look forward to the final study results.

REVIEWER	Arie Hasman I am a retired professor in Medical Informatics.
REVIEW RETURNED	28-Jan-2017

GENERAL COMMENTS	The manuscript is well written. I have only a few remarks. It is estimated that 2% of the patient encounters that are not exposed to the MayoExpertAdvisor will have a change in the care plan that would have been suggested by information directly reported by the MayoExpertAdvisor system had it been available; and that MayoExpertAdvisor will increase the percentage of patient encounters with a change in their care plan to 4% (relative risk = 2). This change is considered the minimum clinical relevant effect. To me a 4% change can hardly be considered relevant for the use of an advisor system. The authors do not state whether the same patient can be encountered more than once in the study. If that would be possible, how does that effect the statistics?

VERSION 1 – AUTHOR RESPONSE

RESPONSE TO REVIEWER 1

Thank you for your careful review and constructive criticism.

Critique 1: I find the lack of references to similar studies of guideline adherence in the literature to be unacceptable. Many studies have been done over the last 20 years very similar to this one. Failure to recognize past attempts does not make your study innovative. Here are a few examples just to get you started. I think any new trial of an old technology should carefully review what has been tried, what worked, what didn't, and then there should be an explanation of why the current proposed trial has a chance to succeed.

Response: The introduction has been revised to include references to similar studies and a discussion of the factors that may have influenced the success of previous interventions has been added. We added an explanation as to why we think our intervention may be more successful than previous interventions.

Critique 2: I think you should present the baseline adherence rate to these guidelines in this paper. I can't believe that you haven't already done that, so why not report it?

Response: Baseline adherence was measured, but will be reported with the main results of the trial.

Critique 3: Why didn't you stratify the practices based on baseline adherence to the guidelines? What if the intervention group has a higher baseline rate of adherence and therefore does not need the advisor? This could make your study fail, due to adverse patient/physician/practice selection.

Response: We appreciate your input. It is too late for us to incorporate this into the design as the trial began in June.

Critique 4: I think you should be measuring and reporting the percentage of patient visits in which the advisor was available and that the clinician actually looked at it. In my experience, systems like this have low utilization, especially after the initial excitement has worn off, which it most likely will after 3 months or so and then you will have 3 more months of data collection. I would still expect you to analyze your data with an "intention to treat" model, meaning that you must still include all the patients/clinics even if the physicians didn't use the system.

Response: Use of the tool is being measured using log file data. Information on this is somewhat hidden within the data sources section of the paper and thus we have revised the secondary outcomes section to include an outcome looking at use of the tool. Hopefully this will clarify to the reader that we will be measuring and reporting use of the tool.

Critique 5: I think you should also make several "outcome" measures on the patients to help ensure that the new guidelines don't inadvertently lead to increased patient visits, ED visits, hospitalizations, or deaths. While I believe the guidelines are good, if there is an error in any part of this system and the clinicians follow these guidelines you may have problems. I would hope to see some improvement in patient health as a result of following these guidelines, but with such a low expectation of improvement in guideline adherence (see below), this is certainly unrealistic. You should at least look for these changes.

Response: We appreciated your comments on measuring downstream effects. It is too late for us to incorporate this into the design of this trial, but if our intervention is successful a subsequent study where we look for changes in ED visits, hospitalizations, cardiovascular outcomes, and deaths, especially for the patients with atrial fibrillation and heart failure patients is essential and will be undertaken.

Critique 6: I find it hard to believe that a study of this complexity is worth doing for only a change in adherence from 2% to 4% (see page 8 line 32). While you may get statistical significance, that change does not seem clinically or practically significant at all. Unless, your baseline compliance is already over 75%. And if it is over 75%, then I would also question why one would put so much effort into a study for such a small benefit. Please try to explain, why you think an organization should expend so much time and energy for such a small improvement, when there are major problems with quality in our healthcare
system.

Response: We chose 2% based on the median improvement in prescribing in a systematic review of similar studies and have included this information in the statistical consideration section. We agree that this is the absolute minimum improvement we would consider clinically relevant, and hope to see a larger effect in the trial.

Critique 7: Page 9 line 8 HIPAA is spelled wrong!

Response: We have changed the misspelled HIPAA, thank you for your careful reading!

RESPONSE TO REVIEWER 2

Thank you for your careful review and constructive criticism.

Critique 1: Most of my concerns have to with matters that need to be more fully addressed in the manuscript. The first problem has to do with the rationale. The authors state that “many interventions have been designed to address these gaps between evidence and practice, yet most achieve only minimal gains in guideline adherence”. What is it about those interventions that have yielded only minimal gains? What is wrong with the “received” or conventional approach to decision support or computer-based clinical guidelines? Why is there an expectation that the use of MayoExpertAdvisor would yield superior results? What differentiates it from the other tools/approaches? It would be worth developing this argument in a paragraph or so.

Response: We have included a paragraph in the introduction that details the factors that may have influenced the success of previous interventions and added an explanation as to why we think our intervention is different from the other tools that have been studied.

Critique 2: It would be useful to know more about the Level 3 and 4 validation efforts. What principles or findings guided the iterative redesign effort? It would also be useful to know a little more about the results of the pilot beyond general satisfaction of those who were surveyed. For example, what was deemed to be of value and what needed to be revised. How was the system designed to be integrated with clinical workflow?

Response: The Validation of MayoExpertAdvisor section has been revised to include more information about level 3 and 4 validation efforts as requested.

Critique 3: Why will physician assistants be included in the study and residents excluded?

Response: Residents were excluded as the oversight committee for residents participating in research was not enthusiastic about residents participating in the study. All other clinician types who provided primary care for patients were included.

Critique 4: In a sentence, please explain how the system employs NLP to extract data that would be otherwise unavailable. Similarly, a couple of sentences on the design of the user interface would be informative.

Response: The MayoExpertAdvisor Section has been revised to include more information about the use of NLP and the user interface.

Critique 5: Is it possible to use log file data in lieu or in addition to the survey to derive efficiency measures in managing the 3 clinical conditions or use of risk calculators? This would have greater validity than self-report measures.

Response: Log File data will be used in addition to survey data to measure use of MayoExpertAdvisor. This was somewhat hidden within the data sources section of the paper and thus we have revised the secondary outcomes to include use of the tool as a secondary outcome.

Critique 6: Interviews with a small number of the study participants would be illuminating. Of course, the choice of methods is at the discretion of the authors.

Response: We are currently planning a qualitative study with focus groups including both frequent and infrequent users of MEA. We agree this will be very informative.

Critique 7: The guidelines workflow diagrams in appendices 2 and 3 are very hard to read. There is ample room to expand them on a page.

Response: We revised the diagrams, but when re-saved in the acceptable format for submission, they were no more readable than the original appendices. Thus, we kept the original. Since they are very legible in the web format, which is how they will be available with the article, we are hopeful this will not be an issue for readers.

Critique 8: The authors appropriately raise the limitation regarding generalizability given the use of proprietary systems only available to the Mayo Clinic. A follow-up sentence (or so) should articulate how this research can still make a substantive contribution to knowledge regarding clinical practice (or decision support). In other words, what will be the nature of the take home message to those who wish to achieve similar clinical goals related to improved adherence to guidelines but lack access to MayoExpertAdvisor.

Response: The limitations section was revised to better address the potential for broad applicability of our study findings.

RESPONSE TO REVIEWER 3

Thank you for your comments.

Critique 1: This reviewer is not a specialist is statistics so the appropriateness of the precise study size will need to be determined by another reviewer.

Response: We are confident the sample size calculation provided by our study statistician is accurate.

RESPONE TO REVIEWER 4
Thank you for your comments. We also look forward to the final study results.

RESPONSE TO REVIEWER 5

Thank you for your careful review and constructive criticism.

Critique 1: It is estimated that 2% of the patient encounters that are not exposed to the MayoExpertAdvisor will have a change in the care plan that would have been suggested by information directly reported by the MayoExpertAdvisor system had it been available; and that MayoExpertAdvisor will increase the percentage of patient encounters with a change in their care plan to 4% (relative risk = 2). This change is considered the minimum clinical relevant effect. To me a 4% change can hardly be considered relevant for the use of an advisor system.

Response: We chose 2% based on the median improvement in prescribing in a systematic review of similar studies and have included this information in the statistical consideration section. We agree that that this is the absolute minimum clinically relevant effect.

Critique 2: The authors do not state whether the same patient can be encountered more than once in the study. If that would be possible, how does that effect the statistics?

Response: The reviewer raises an important point about the possible variance inflation correction that may need to be applied if a patient has multiple encounters in the study. While ideally, we would like to account for multiple encounters, there are several pragmatic reasons why this may not be feasible. First, there are already higher levels of clustering being incorporated into the analysis and the event rate is expected to be low. With multiple levels, in the hierarchical linear model sense, convergence issues may occur (essentially, there may be insufficient variability in some levels of clustering to estimate the random effect). Next, with the de-identification of the patient records, per IRB requirement, the analysts may not be aware of the clustering within patient. Finally, given the time window for enrollment, multiple encounters are expected to be minimal. The analysis plan in the paper has been amended to acknowledge these limitations.

VERSION 2 – REVIEW

REVIEWER	Dean F. Sittig School of Biomedical Informatics, University of Texas Health Science Center at Houston, TX
REVIEW RETURNED	10-Mar-2017

GENERAL COMMENTS	I am disappointed that the authors chose not to address my main concerns regarding this study.

REVIEWER	David Kaufman Biomedical Informatics, Arizona State University
REVIEW RETURNED	19-Apr-2017

GENERAL COMMENTS	The authors have been largely responsive to the reviewers' concerns. They have addressed most comments to the extent possible. This is a very solid proposal that is both timely and of general interest. The manuscript is clearly written, the rationale is solid and the study design is well presented. At the same time, the limitations of this work more clearly come across in both the manuscript and in the response to the reviewers' letter. I only have a few minor comments. They all concern less than fully adequate responses to key issues: The revised paper does not adequately address the user-centered design process or the usability studies. Clinicians worked with graphic designers to develop the user interface. A user-centered design process in which the team researched the needs of clinicians and aligned the design of the application to meet those needs was used. Industry-best practice design principles for usability23 were used throughout the designprocess. The description is couched in very general terms and is not informative. A sentence or 2 of specifics would suffice. For example, one can use the screenshots to comment on the layout or arrangement of screen items. Just an aside, the contrast of light grey font against a blue background (at least as it appears in the image) is problematic (not a problem for the paper, per se). For the testing of the interface (Level 3), please report a couple of findings to provide us with a little more context on the process informed the iterative design. The last sentence in the paper (see below) makes a rather obvious point. The paper describes a clearly articulated model and approach to decision support and guideline-based prescribing practices. Surely, there are more compelling implications that can be draw from this work. The approach should be generalizable even to settings that lack the specific Mayo resources. Similarly, the findings from the study However, the general goal of improving guideline based prescribing practices for common chronic diseases through a clinician decision support tool, designed and implemented with iterative clinician input, is broadly applicable.

REVIEWER	Arie Hasman Dept. of Medical Informatics, Academic Medical Center, University of Amsterdam, the Netherlands
REVIEW RETURNED	13-Mar-2017

GENERAL COMMENTS	On page 4 the authors several times use the term intuitional guidelines (also in the table 1 caption). I assume that institutional guidelines are meant. The reviewers have made some suggestions where the answer of the authors was that it was too late to take them into account. I ask myself why the facts that gave rise to the suggestions cannot be seen as limitations of the study and therefore should be reported in the limitations section. The de-identification process could have used pseudonyms making it possible to follow individual patients without knowing their identity. I understand that the chance that a patient has more than one visit during the study period is rather small, so pseudonymisation will not be very effective.

VERSION 2 – AUTHOR RESPONSE

REVIEWER 1

Critique 1: I am disappointed that the authors chose not to address my main concerns regarding this study

Response 1: It is not entirely clear what this reviewer is referring to. As an itemized response was provided to all previous critiques, I can only assume that the concern was that we did not change the study design. We have listed some of the changes recommended that we were not able to implement as the study was already underway as limitations. This seems like it would be an adequate response to reviewer number 1 concerns. A change in study design when then trial was already underway was not possible.

REVIEWER 5

Critique 1: On page 4 the authors several times use the term intuitional guidelines (also in the table 1 caption). I assume that institutional guidelines are meant.

Response 1: This typo was corrected

Critique 2: The reviewers have made some suggestions where the answer of the authors was that it was too late to take them into account. I ask myself why the facts that gave rise to the suggestions cannot be seen as limitations of the study and therefore should be reported in the limitations section.

Response 2: The authors added revised the limitations section to include the limitations in study design noted by other reviewers.

Critique 3: The de-identification process could have used pseudonyms making it possible to follow individual patients without knowing their identity. I understand that the chance that a patient has more than one visit during the study period is rather small, so pseudonymisation will not be very effective.

Response 3: I am not sure what the implications of this comment are. The initial comment seems to recommend pseudonymisation but then states that this would not be effective. No action was taken on this critique.

REVIEWER 2
Critique 1: The revised paper does not adequately address the user-centered design process or the usability studies. A sentence or 2 of specifics would suffice. For example, one can use the screenshots to comment on the layout or arrangement of screen items. Just an aside, the contrast of light grey font against a blue background (at least as it appears in the image) is problematic (not a problem for the paper, per se).

Response 1: We added a few sentences to elaborate on the user-centered design process.

Critique 2: For the testing of the interface (Level 3), please report a couple of findings to provide us with a little more context on the process informed the iterative design.

Response 2: We added a few sentences to elaborate on the process informed the iterative design.

Critique 3: The last sentence in the paper makes a rather obvious point. The paper describes a clearly articulated model and approach to decision support and guideline-based prescribing practices. Surely, there are more compelling implications that can be draw from this work. The approach should be generalizable even to settings that lack the specific Mayo resources. Similarly, the finding from the study.

Response 3: This sentence was removed as it was highlighting an obvious point. Additional limitations were added to this section, as per the suggestions of other reveiwers.

VERSION 3– REVIEW

REVIEWER	David Kaufman Arizona State University, United States I have no competing interests. In interest of full disclosure, I am a Mayo Clinic Research Affiliate. However, I do not collaborate or work with any of the authors.
REVIEW RETURNED	16-Oct-2017

GENERAL COMMENTS	This is a well designed study that purports to test a decision support tool for improving guideline adherence across conditions. The study is well motivated with a very sound rationale. MayoExpertAdvisor was a clearly conceptualized and developed with appropriate design considerations. The tool has gone through extensive iterative formative testing and appears to be ready for a larger scale test of its efficacy as described in the proposal. The research design is meticulously laid out. This is clearly a very promising work. Once again, the authors were largely responsive to the reviewers' comments. The timeframe issue remains somewhat unusual given that this is a research protocol proposal and the study has been underway for quite some time. In addition, the paper makes reference to a study that began June 1, 2016 and is expected to run for 6 months. In addition, both future and present tense is used to refer to the state of the research. This matter needs to be reconciled. If it's a study in progress, then the manuscript should reflect that.