Article details: 2016-0031	
Title	The PR1MaC Study: a pragmatic randomised trial of the integration of chronic disease prevention and management services into primary care
Authors	Martin Fortin MD MSc, Maud-Christine Chouinard RN PhD, Marie-France Dubois PhD, Martin Bélanger MD, José Almirall MD PhD, Tarek Bouhali MD, Maxime Sasseville RN
Reviewer 1	Michael E. Green
Institution	Departments of Family Medicine and Public Health Sciences, Queen's University, Kingston, Ont.
General comments (author response in bold)	Specific comments: Comment: 1. The title only reflects one of the three methods used to evaluate the intervention. This could be revised to be more inclusive of all 3 of the study components. **Response: We tried to include in the title all 3 of the study components, but all attempts ended up in long titles. We decided to keep only the most important study design in the title.** Comment: 2. In "Setting and Subjects", add "in Quebec" after "primary care physicians" in the second sentence to clarify that these models are specific to that province. **Response: The clarification was added to the text.** Comment: 3. Table 1 seems to be an appreciated version of the table used in the BMC HSR paper from 2013. I read that paper and found the original table very helpful. Consider simply re-using that table. The additional sections on training of the CPDM providers and follow up etc. were helpful additions. **Response: We own the copyright of the BMC HSR paper from 2013, and reproduced the table acknowledging where it was published first.** Comment: 4. I would suggest introducing and referencing figure 1 during the section on Study Design. I found it helpful to have up when I was reading this section. Consider adding appropriate labels to this figure to show which groups are being compared when in each of the 3 study designs. Ie. Add labels indicating that the pragmatic RCT component compares groups A and B at 3 months, that the before after is group A only at baseline and 12 months and that the quasi-experimental design is both A and B vs C. **Response: We now introduce figure 1 during the section Study Design. We added appropriate labels to figure 1 to show which groups are being compared in each of the 3 study designs.** Comment: 5. Methods "Pragmatic Randomized Trial". The first sentence does not flow well. Consider switching the order to "Participants reported sociodemographic data and completed and initial set of questionnaires at baseline" **Response: We corrected the text with the sentence suggested by the reviewer.** Comment: 6. In the section on "Variables and Outcome Measures" and "sample size" Please define what was used to represent a "meaningful change" in the heiQ domains (from the standpoint of the researchers or clinicians in terms of how much of a difference matters), and also what effect size was used for the power calculations (this was in the methods paper but seems to have been cut here). **Response: We added the effect size used in the section 'Sample size and statistical power'. The measure of "meaningful change" that we used was suggested by the developers of the heiQ. However, there is not a straightforward unit to explain how much a difference matters. The simplest definition that we can provide of a 'meaningful' change is that each individual difference in score on a domain is meaningful (reliable) if the different exceeds the 1.65 threshold when it is divided (corrected) by the standard error of the difference. We added this explanation to the text.** Comment: 7. In the interpretation section I think that more explicit reference to what the observed effects sizes might mean clinically would be helpful. Most clinicians do not usually use Cohen's d as a measure of effect size regularly and will want some guidance as to what this means for the expected impact of the intervention on their patients. They would be more used to seeing results such as this expressed as adjusted mean differences in

	the actual scores and will want to know in plainer language how much of an improvement was found and what proportion of patients would expect to benefit. I note for example, that some the statistically significant differences in Table 4 have very small Cohen's d even if they are statistically significantly different (ie. SF-6D, PCS). **Response: As we explained in a previous response, we interpret from the reviewers' comments that Cohen's d as a measure of effect size does not mean much for clinicians who do not use it regularly. For this reason, we have calculated adjusted mean differences in the actual scores and their 95% confidence interval for comparisons between groups.** Comment: 8. In tables 5 and 7 the labels should be clearer. PD and DP are used to mean the same thing in different tables and it might be better if there is room to write out longer abbreviations or titles. **Response: We corrected the use of PD and used longer abbreviations in both tables. Unfortunately, there is no room in the tables for longer titles.**
Reviewer 2	Wilson Pace
Institution	University of Colorado, Family Medicine, Aurora, Colo.
General comments (author response in bold)	Comment: 1. No analysis of the differences at baseline between the groups is provided. The reader is left to infer the similarity of the two groups by review of the demographic data. Indicating which if any variables were significantly different would be helpful. As two variables were controlled for in the regression analyses one would assume there were differences in these variables at baseline though this is not evident from a review of Table 2. **Response: Please, see higher above in this text our response to a similar comment. The variables at 3 months were adjusted for their baseline values in the regression analysis. In the previous version of the manuscript, it was written in the footnote of Table 7 that analyses were controlled for age and sex, but this was corrected.** Comment: The protocol called for 3 visits per patient with the behavioral change specialists, whether more visits were possible is not clear from the methods section. The average number of visits appears to be inferred at 2.4 instead of directly measured. The range of visits is not provided (which may be mute if 3 visits was the maximum allowed but again this is not clear) nor is it clear if the 2.4 visits refers to an average of both Groups A and B or just to group A. **Response: The average number of visits was not inferred, it was measured. It refers to an average in group A during the randomized trial. This was included in the text. We also included the range, 0 to 5 encounters, because there were patients who visited more than one provider.** Comment: The short time frame for follow up for the true RCT component of the study is a major concern as short term, self-reported behavioral change during an educational intervention is common if not almost universally seen in studies of this type. No longer range follow up was possible due to the immediate provision of the intervention to the 'control' group. The possibility of a desirability bias at three months is much higher than the authors seem to acknowledge. Though this is somewhat minimized by the 12 month follow up it most likely heavily contaminants the three month data. Even a 3 month further delay in the provision of the intervention with an outcome assessment between groups at 6 months would have greatly strengthened the results. **Response: We added to the text the acknowledgment that the short time frame for follow up for the RCT component of the study is the main limitation of the study.** Comment: The quasi-experimental follow-up is helpful in determining long term effectiveness of the intervention, but the differences drop to only two measurements and the design is considerably weaker than a truly randomized trial. The use of trained professionals in a wide variety of disciplines makes the intervention less amenable to dissemination in to clinical practice. It is not clear what was 'transferred' to the primary care physicians for follow up since the group received no training as to how to integrate with the intervention. **Response: Physicians' follow up did not mean to continue with the intervention but just to continue monitoring their patients.** Comment: 2. Figures 2 and 3 are not interpretable with the current legends. Do they refer to just group A vs. group B or to groups A and B relative to group C? **Response: We have re-written the legends and hope that they are clearer now.**

	Comment: 3. While this study demonstrated a number of significant differences at three months the lack of longer term follow up in the truly randomized groups and the loss of almost all of the outcomes by 12 months in the quasi experimental component of the study significantly weakens the findings. The intervention appears to be difficult to sustain in routine practice and the minimal outcomes observed are not likely to sway policy makers to embrace the concept. It is unclear that these findings are a significant advance in primary care behavioral change for multi-morbid individuals. **Response: We acknowledge that there is room for improving the intervention. Decision-makers have been already interested in the results of the study because it represents a step forward towards implementation of interdisciplinary practice.**
Reviewer 3	Nicola Vanacore
Institution	National Institute of Health, National Center for Epidemiology, Rome, Italy
General comments (author response in bold)	Comment: 1. The authors should clarify the sample size in the definition of the percentage of patients improving on the heiQ in two groups. **Response: This information was included in the Appendix.** Comment: 2. The 144 patients received the intervention but refused to participate in the research. What were the reasons? How this may influence the external validity of the findings. The authors should discuss this important issue. **Response: We respected patients' decision to refuse to participate and did not demand the reasons for not participating. Therefore, we could not collect any information from them. The impact of this on the composition of the sample of participants is not known. We discuss this situation in the limitations of the study.** Comment: 3. The lost to follow-up at 3 and 12 months (n = 54 in total) should be considered and discussed in relation to findings. **Response: We considered the lost to follow-up and added a few sentences about this to the discussion.** Comment: 4. In my opinion, the authors should specify how the FRQS peer reviewed the protocol and manages the funding. In particular how the modality of funding may have influenced the conduction of the study. **Response: The modality of funding did not have any influence in the conduction of the study.**