Author's response to reviews

Title: General practitioners' views and preferences about quality improvement feedback in preventive care: a cross-sectional study in Switzerland and France

Authors:
Paul Sebo (paulsebo@hotmail.com)
Hubert Maisonneuve (hubert.maisonneuve@unige.ch)
Jean-Pascal Fournier (jean-pascal.fournier@univ-nantes.fr)
Nicolas Senn (nicolas.senn@chuv.ch)
Dagmar Haller (dagmar.haller-hester@unige.ch)

Version: 1 Date: 24 Jun 2017

Author’s response to reviews:

IMPS-D-17-00118

General practitioners' views and preferences on quality improvement feedback in preventive care: a cross-sectional study in Switzerland and France

Paul Sebo; Hubert Maisonneuve; Jean-Pascal Fournier; Nicolas Senn; Dagmar Haller

Implementation Science

Reviewer #1:

I thank the authors for their manuscript. The paper presents a questionnaire study among Swiss and French GPs exploring their perceived usefulness toward receiving feedback on their preventive care performance, and how this feedback should be provided. The main results are that 44% found feedback useful; younger and those doing reportedly more preventive care were more likely to find feedback useful; its main use would be 'knowing about study results' and 'improving practice'; and the preferred feedback type would be 'a brief report' or 'a report with specific information about best practice', but definitely not 'face-to-face discussion with study investigators'.

The study is interesting because it explores GP preferences for receiving feedback before any feedback intervention is developed; the results may therefore guide future feedback designs and increase feedback acceptance in this setting. The study has a high response rate (47%). However, in my opinion there is a number of concerns that need to be addressed before publication.
1. The introduction of the study is lacking some clinical background regarding prevention in primary care e.g. why prevention is important or complex, why there may currently be suboptimal performance (and perhaps report some prevention guideline adherence rates in other settings), and why feedback could be effective to improve this. Now, it is not completely clear to the reader what certain preventive measures e.g. "blood pressure, weight and height measurement" may exactly prevent and how they may be beneficial for the patient. The study also lacks information about key differences between France and Switzerland regarding preventive care and current state of guideline adherence/implementation. This information may be essential to understand potential differences in outcome measures between the countries.

- Why prevention is important: we completed the introduction section: “Preventive care is an important part of GPs’ tasks which potentially contributes to reducing the burden of chronic diseases, such as heart diseases or cancers, and has the potential to decrease medical costs”. (ref: for references see manuscript/clean version)

- Why there may be suboptimal performance: we completed the introduction section: “GPs experience difficulties adhering to guidelines because of organizational constraints (in particular lack of time), insufficient financial compensation for providing preventive care, lack of awareness or familiarity with the guidelines, absence of agreement between the various guidelines, and difficulties in applying some recommendations in daily practice”.

- Why feedback could be effective to improve adherence: we acknowledge both reviewers’ experience in exploring this question. We slightly modified the introduction section to state more clearly why feedback is expected to improve adherence: “Several authors have shown feedback to be useful to improve health care, including preventive care. (ref) It is based on the belief that healthcare professionals improve their practice when they receive feedback following audit showing suboptimal performances. (ref) Using behavior changes theories, it is hypothesized that feedback may change GPs’ awareness and beliefs about current practice, change their perceived social norms, and/or may lead them to focus on sub-goals. Within a framework based on control theory, GPs’ self-assessment of clinical performance and targets following audit and feedback serve to mobilize their intentions to improve their practice and adherence to guidelines. (ref Gude) A range of factors both specific to each physician (emotions, core values…) and linked to the environment (resources, workload…) determine whether physicians develop intentions to change in response to feedback, and manage to put intentions into practice. (ref Payne & Hysong)

- What certain preventive measures e.g. "blood pressure, weight and height measurement" may exactly prevent: we completed the introduction section: “Prevention is particularly important for the management of major modifiable risk factors, such as smoking, dyslipidemia, obesity and high blood pressure. Preventive care is an important part of GPs’ tasks which potentially contributes to reducing the burden of chronic diseases, such as heart diseases or cancers, and has the potential to decrease medical costs. (ref) For example, measuring blood pressure within the practice is a noninvasive and inexpensive approach which may play a crucial role in the early detection (and treatment) of patients at high risk for the development of cardiovascular disease”.

- Key differences between France and Switzerland regarding preventive care and current state of guideline adherence/implementation: we completed the introduction section: “In Switzerland, a national program named EviPrev was recently launched to develop local guidelines, but these recommendations have not yet been implemented in clinical practice. (ref) In France, the medical authorities decided to focus their guidelines on the management of diseases rather than on preventive care alone. Though some aspects of preventive care are addressed in disease-specific guidelines, French GPs meet more difficulties accessing national preventive recommendations”.

“The difficulty of adoption and transfer was confirmed by several studies carried out in the US, which showed that the rates of preventive care were suboptimal: overall, only one-half of recommended preventive services were usually provided. In contrast, we previously reported high adherence by Swiss and French GPs to most recommendations for prevention (>70%), though certain measures were less often provided (above all, annual influenza vaccination for at risk-patients <65 years old). (ref) We also showed that, compared to French GPs, those practicing in Switzerland tended to provide slightly more measures.”

2. In the methods section there are many things unclear about the questions in the questionnaire. I do not understand the use of feedback "to know study results" and "to follow-up what is done in their practice", nor what the various types of feedback would entail exactly, such as "brief report and individual results". The authors need to clarify what is meant and state how these questions were literally (translated) formulated. The approach that the reasons for and type of feedback were multiple choice questions and that multiple answers were possible is a surprising. This would invite people to either say 'yes' to all, or select only some reasons/types even though the other reasons/types may just have been slightly less preferred. Lugtenberg et al (reference 16 in the paper) used a Likert scale which seems a more logical approach. The authors state that the selection of the feedback interventions was consensus-based following literature review, but there were still many options missing and there was no "other" option to get new ideas from the respondents. For example, the 2012 Cochrane review of feedback interventions (reference 15 in the paper) reports many types of feedback delivery including the format (verbal/written/electronic) and source (supervisor/colleague/investigator), and many co-interventions that can be used with it.

In this study, as stated in the methods section, GPs were asked:

i) to estimate the usefulness of a feedback regarding their preventive care practice (indispensable, very useful, rather useful, not very useful, not useful),

ii) to explain for which reason(s) they would find a feedback useful (multiple answers allowed: to know the study results, to compare themselves with colleagues, to modify or improve their practice, to use the feedback to follow-up what is done in their practice),

iii) which type of feedback they would like (multiple answers allowed: a brief report, a brief report and individual results, detailed results regarding their practice compared with the study results, a report and specific information regarding prevention best practice, a contact with the study investigators to discuss the results, a local quality circle meeting to discuss the results).
The selection of the feedback interventions was based on a consensus within the study team following a review of the literature.

In relation to the reasons why GPs would find a feedback useful, we understand the reviewers’ remark. In translating French into English, the meaning of “to know the study results” and “to use the feedback to follow-up what is done in their practice” became probably unclear for the reader.

We changed “to know the study results” to “to know GPs’ overall performance”. We also changed “to use the feedback to follow-up what is done in their practice” to “to use regular feedback interventions to follow-up what is done in their practice”.

Note that the questionnaire was developed in French, translated into English only for publication purposes, and it was pretested with seven community-based GPs to ensure that the questions were understandable. There were no comments on these questions during the pretest.

In relation to the type of feedback, “a brief report” means a general report summarizing the goal and the importance of feedback but without giving the GP’s performance, whereas “a brief report and individual results” means reporting in addition the GP’s performance.

We changed “a brief report” to “a brief report, i.e. a report providing general information about GPs’ overall performance without individual and detailed results”.

We accepted multiple answers for two questions (the reasons for and type of feedback), because we think that GPs could find a feedback useful for several reasons and/or they would like to receive several types of feedback. “Other reason” and “other type of feedback” existed in the list as possible responses. Yet these response options were not selected by GPs in our sample and therefore did not appear in our original results. We have now added them.

We did not use Likert scale for these two questions, but we agree with the reviewer, this could have been an interesting option for asking these questions.

The reviewer is right; there were other types of feedback and other types of format that we could have selected for these two questions. After reviewing the literature (see ref. in the methods section) we decided to keep the five options for the first question, respectively the seven for the second question, that seemed to us the most interesting, important or pertinent to study. As stated above, “other” was one of the proposed answers but was not selected by the responders. In designing our survey, we placed a priority on developing as short a questionnaire as possible in order to reduce the burden of participation for GPs. With questions offering many response options, we were afraid that certain questions would have been completed incorrectly, for example at random.

We completed the limitations section: “After reviewing the literature we decided to keep only five response options for the question assessing the reason(s) why GPs would find feedback useful, respectively seven options for the question about the type(s) of feedback they would like. These options seemed to be the most interesting, important or pertinent to study in the context of
primary care. We decided to limit these response options, because with questions offering many possible response options, we were afraid that certain questions would have been completed incorrectly, for example at random. Though this restricted selection could result in a certain degree of information bias, we do not believe that this was in fact the case. Indeed, though “other reason” and “other type of feedback” were response options proposed to responders, they were never selected by GPs in our sample. In addition, no GP (n=7) suggested in the pretest phase that we add other response options to the list.

3. The way the variables are handled in the statistical analysis seems arbitrary. There are many continuous variables that are categorized, and it is unclear why. For example, in Table 4 the variable 'number of prevention measures carried out' was dichotomized into less than 10 versus equal to or greater than 10. Age categories were chosen 0-35, 35-44, 45-54, 55-64 and 65+; it seems unclear why these categories have been chosen and not handled as continuous variable. Not handling e.g. age as continuous variable is throwing away potentially important information; therefore the authors should reconsider how they handle variables or clearly state their reasons for not doing so in the paper. Furthermore the abstract states nothing about any statistical analysis.

Age was recorded as a categorical variable. We have an increasing tendency to record age as categorical variable in studies recruiting doctors, in particular GPs. Indeed, some GPs might think that their identity could be known if their exact age was recorded. However, as the variable was categorized into 5 categories, not only two, we believe that this variable provides us with sufficient information. We completed the methods section: “age group (<35, 35-44, 45-54, 55-64, >64)”

The variable “number of preventive measures” was dichotomized into < or >=10 for simplification purposes in Table 4. This binary variable was only used in the multivariate analyse to assess if, among various other factors, those carrying more measures of preventive care were more likely (or not) to consider feedback useful. Our goal was not to discuss in detail the results of each group of GPs according to the number of preventive measures provided. As stated in the methods section, we used this cut-off because this is approximately the median of the distribution: “10 was the third quartile”.

As rightly suggested by the reviewer, we completed the abstract: “Chi-square tests were used to compare frequencies. Multivariate logistic regression was used to identify factors associated with GPs considering feedback as useful.”

4. It is misleading that the results section and abstract state that the study was conducted among 1100 GPs, whereas it was actually conducted among 518 GPs (response rate 47%). Next, if 1100 GPs were invited and 47% responded, it may be interesting to see what are the differences between responders and non-responders. Given that the GPs were invited by post, the authors may have some information besides the address (e.g. gender and age) of the non-responders.

We agree with the reviewer’s remark, the abstract is indeed unclear. We modified the abstract: “Methods: This cross-sectional study was conducted in 2015 among randomly selected
community-based GPs in two regions of Switzerland and France” and “Results: 518 of 1100 GPs (47.1%)”

By contrast, in the results section, it is already stated that “among the 1100 GPs who were contacted at random, 518 (47.1%) participated in the study”.

Unfortunately, we cannot assess the differences between responders and non-responders, because we did not collect any data on non-responders. We added this limitation in the limitations section: “As we did not collect any data on non-responders, we could not assess potential differences between responders and non-responders”.

However, our sample appears to be relatively representative of all community-based GPs practising in Switzerland and France (data from Pays de la Loire). We completed the results section: “Our sample appears to be relatively similar in age and gender to all community-based GPs practising in Switzerland (professional organisation of Swiss physicians, 2016: median age 54 years (vs. 54 years in our study); men 59% (vs 61% in our study)) (ref) and France (Pays de Loire, 2013: median age 51 years (vs. 56 years in our study); men 57% (vs. 66% in our study) (ref)” and the limitations section: “However, our sample appears to be relatively representative in terms of age and gender of all community-based GPs practising in Switzerland and France (data from Pays de la Loire)”.

5. The numbers in the results table do not seem to add up. For example, Table 1 says that there were 160 respondents from France and 350 from Switzerland (adding up to 510) whereas the paper states that there were 518 respondents. Table 3 states that there were 163 French and 355 Swiss respondents (=518), Table 2 states 250 were under 55 years and 266 55+ (=516), 318 males and 191 females (=509), and Table 2 also reports a total being 718. Furthermore, nothing is stated about the number of missing values and how this may have introduced bias into the study results.

We know that 163 French and 355 Swiss GPs participated in the study (see the results section and/or Table 3). However, only 160 French and 350 Swiss GPs (i.e. n=510) are reported in Table 1, because 8 GPs did not answer the question regarding the usefulness of a feedback.

We added as footnote in Table 1 and 2 “numbers do not add to 518 because of missing data”.

We believe that the risk of bias due to missing data is probably very small, because there were only few missing data in our study. We completed the limitations section: “We believe that the risk of bias due to missing data is probably very small, because there were only few missing data in our study”

Sorry for “n=718” in Table 2. It was a typo. We changed “718” to “518”.

6. Half of Table 1’s contents is unnecessary: the perceived usefulness is dichotomized so if we know 41.4% of men found feedback useful, we already know that 58.6% found it not useful. The 58.6% need not be reported. Furthermore Table 1 and 4 overlap in the message they are trying to send, namely what are the determinants for finding feedback useful or not. Table 1 may be
clearer if it just reports baseline characteristics of the respondents without stating whether or not they perceive feedback as being useful, i.e. % males, mean age, % from France, mean working days per week, etc. The percentages of respondents finding feedback useful can easily be merged into Table 4 into a new column.

We agree with the reviewer. We modified Table 1 (baseline characteristics) and the title: “Table 1. GPs’ socio-demographic characteristics (N=518)”.

We also removed “%” at the end of each categorical value because “%” is already noted at the top of the table.

7. The discussion lacks a clear structure. I suggest using headings in accordance with the widely accepted structure described here: http://www.bmj.com/content/318/7193/1224. Next, the authors state (p13 lines 159-163) that feedback is more likely accepted if feedback is provided written and not face-to-face with clinicians. It would be useful if this statement is assessed in the light of Ivers’ Cochrane review (reference 15 in the paper) which also shows that feedback + outreach visits (= face to face by investigators) are more effective than only feedback. Further, the limitations lack a critical appraisal of selection/response bias and the potential bias induced by the limited number of answer categories in the questionnaire (see comment 2). Due to these biases the results may have been overinterpreted. Finally, the authors state in the discussion (p14 line 173-177) that quality circles should be implemented whereas this clashes with their results that "face-to-face..........discussion with study investigators" is highly not preferred. I do agree with the authors that PDCA cycles are beneficial, but the authors seem to suggest these can only conducted with investigators in a quality circle whereas they can also be conducted by GPs themselves (potentially supported by a feedback system e.g. a quality dashboard) without investigator involvement.

We thank the reviewer for this interesting reference. In this paper, the author suggests the following structure for the discussion section of scientific studies:

• Statement of principal findings
• Strengths and weaknesses of the study
• Strengths and weaknesses in relation to other studies, discussing particularly any differences in results
• Meaning of the study: possible mechanisms and implications for clinicians or policymakers
• Unanswered questions and future research

The structure we used to guide our discussion is broadly similar, though based on another reference: summary of main findings, comparison with existing literature, limitations (or strengths and limitations), conclusion (or conclusion and implication for practice). See for example Writing and Publishing in Medicine by Edward Huth (third edition), page 85:
“Discussions should briefly recapitulate the answer that comes from the findings and the strength of that answer; they should also present supporting evidence from previous research and deal with counterevidence. Discussions should close with assessments of generalizability of the findings and with implications for further research”.

We added subheadings in the discussion section in order to better separate the different sub-sections: summary of main results, comparison with the existing literature, limitations, implications for practice and policy, and conclusion.

As suggested by the reviewer, we also completed the discussion section using ref Ivers: “GPs in our study preferred written formats of feedback (brief or detailed report sent to the participants) over verbal feedback (face-to-face contacts with the investigators or local quality circle meetings). Though there may be a perception that feedback delivered face-to-face is more efficient compared to written feedback, the evidence shows little or no difference between the two formats” (ref. Ivers)

We completed the limitations section: “As the answers were self-reported, our findings could partially be explained by the fact that responders may have a natural tendency to over-report positive, socially desirable behaviors (social desirability bias)”. (ref)

We tried to explain above why the questions “reason for feedback” and “type of feedback” had only 5, respectively 6 possible response options. Again, we would like to point out the fact that the option “other” was available for the two items, but not used by the responders.

As rightly suggested by the reviewer, we completed the discussion section regarding PDCA: “PDCA cycles can also be conducted by GPs themselves (potentially supported by a feedback system e.g. a quality dashboard) without investigator involvement”.

8. The paper may improve after being checked by a native English speaker/writer and carefully checked on spelling. For example, in Table 2 and 3 there is a missing "t" in "Contact with the study investigators". Also, "those considering a feedback" (p8 line 103) should be "those considering feedback". There is a big difference between 'useless' and 'not useful'. The tables name 'useless' whereas the main text speaks of 'not useful'.

Sorry for the typo! We changed “contac” to “contact”.

We changed “those considering a feedback” to “those considering feedback”, and “not useful” to “useless”.

We asked a native English GP to check the paper, including the tables.

Reviewer #2:

Thank you for the opportunity to review this paper. The authors aimed to examine general practitioners (GPs)' views on the usefulness of feedback and their preferences regarding how
feedback is provided, through surveys of a random sample of 1100 GPs in Switzerland and France. Results indicated less than 1/2 the sample found feedback to be useful; more favorable perceptions of feedback utility were associated with younger age and greater clinical emphasis on preventive care. GPs reported preferring feedback via a brief report, and a report with specific information regarding prevention best practice, whereas less than 1% would like to discuss the results face-to-face with the study investigators.

The idea of examining practitioner feedback preferences is a sound one, as it improves the likelihood of feedback acceptance and thus the likelihood of feedback effectively changing behavior. The international nature of the sample is a strength of this work. Nevertheless, there are several methodological details that require attention before this manuscript is more publication-ready:

BACKGROUND:

1. The intro spends a long time discussing clinical practice guideline adherence, yet this is a study about feedback preferences. The effectiveness of feedback is well established in health care, including preventive care; I would have liked to see more detailed information about the psychology of feedback perceptions and feedback usefulness; it would have couched the research question far more effectively.

As a prelude to this response, please note that reviewer#1 asked for more information regarding prevention and adherence: “the introduction of the study is lacking some clinical background regarding prevention in primary care e.g. why prevention is important or complex, why there may currently be suboptimal performance (and perhaps report some prevention guideline adherence rates in other settings), and why feedback could be effective to improve this. Now, it is not completely clear to the reader what certain preventive measures e.g. "blood pressure, weight and height measurement" may exactly prevent and how they may be beneficial for the patient. The study also lacks information about key differences between France and Switzerland regarding preventive care and current state of guideline adherence/implementation. This information may be essential to understand potential differences in outcome measures between the countries”.

Therefore, we completed the introduction section according to reviewer#1’s request, adding more information regarding prevention and adherence.

However, we understand reviewer#2’s point of view, and, as a result, we completed the introduction section with information about the psychology of feedback perceptions and feedback usefulness: “Several authors have shown feedback to be useful to improve health care, including preventive care. (ref: for references see manuscript/clean version) It is based on the belief that healthcare professionals improve their practice when they receive feedback following audit showing suboptimal performances. (ref) Using behavior change theories, it is hypothesized that feedback may change GPs’ awareness and beliefs about current practice, change their perceived social norms, and/or may lead them to focus on sub-goals. Within a framework based on control theory, GPs’ self-assessment of clinical performance and targets following audit and
feedback serve to mobilize their intentions to improve their practice and adherence to guidelines. (ref) A range of factors both specific to each physician (emotions, core values…) and linked to the environment (resources, workload…) determine whether physicians develop intentions to change in response to feedback, and manage to put intentions into practice. (ref)"

2. It is not clear what is different about preventive care that would lead the authors to believe that feedback perceptions about preventive care would differ from feedback perceptions about any other type of the care that GPs deliver.

a) In their systematic review, Ivers et al. found that feedback may be highly effective for improving prescribing and less effective for test-ordering or disease management.

Thus, feedback usefulness seems to depend on the outcome tested (prescription, test-ordering, or disease management). It is therefore possible that GPs’ perceptions about preventive care differ from perceptions about other types of care provided.

b) Our aim was to investigate the acceptability of GP practice-based quality improvement feedbacks. Preventive care as a theme was only chosen because the study assessing preventive care practices in Switzerland and France offered the opportunity to find a response to our research question.

We changed in the introduction section “we investigated the feasibility of a practice-based quality improvement feedback for preventive care” to “we investigated the feasibility of a practice-based quality improvement feedback”.

METHOD:

1. Participants were recruited to the study by postal invitation; it is not clear, however, how the questionnaire was administered: in person? by phone? mail? Online?

We completed the methods section: “Each randomly selected GP was contacted by post by a research assistant located in Geneva (for Switzerland) and by the local professional associations Union Régionale des Professionnels de Santé Alsace and Pays de la Loire (for France). GPs were informed about the aim of our study and the practical procedures for completing the questionnaire. The postal letters included a stamped return envelope”.

2. How were the specific preventive practices selected? Along similar lines, the authors state "The selection of the feedback interventions was based on a consensus within the study team following a review of the literature." The resulting list seems rather random; or at least, it does not seem consistent with the domain of available varieties of feedback interventions. This requires more explanation as to how they arrived at the 6 specific forms of feedback delivery in the questionnaire.

In relation to the list of selected preventive practices, we completed the methods section: “They were selected by consensus within the research team following a review of the literature. Ten
preventive practices were included in the study, because they had also been selected by previous authors: blood pressure, weight and height measurements, as well as screening for at-risk drinking (and advice to decrease drinking), for smoking (and advice to stop smoking), and for colon cancer, and, finally, annual influenza immunization for patients >65 years and at-risk patients <65 years). We added two preventive measures to this list: cholesterol measurement, because it is highly recommended, and refraining from systematic screening for prostate cancer, because several medical agencies recently recommended against systematic screening. By contrast, we did not include diabetes screening, because it was considered as targeted screening (i.e. limited to populations with particular risks factors such as obesity) or screening for breast cancer, because this screening is often provided by gynecologists or in the context of screening programs”.

In relation to the choice of feedback interventions: it was selected in a similar way, based on consensus within the research team. The chosen list was pretested with 7 GPs. None of them suggested we should add another intervention to the list. In addition, GPs could select “other type of feedback” in the questionnaire, but none of them did so.

Yet, the reviewer is right; there were other types of feedback that we could have selected. After reviewing the literature (see ref. in the methods section) we decided to keep the seven options that seemed to us the most interesting, important or pertinent to study in the context of primary care, i.e. a brief report (i.e. a report providing general information about GPs’ overall performance without individual and detailed results), a brief report and individual results, detailed results regarding GPs’ practice compared with the study results, a report and specific information regarding prevention best practice, a contact with the study investigators to discuss the results, a local quality circle meeting to discuss the results, and other type of feedback.

We chose to limit the list because if there had been too many response options, we were afraid that certain questions would have been completed incorrectly, for example at random.

We completed the list of possibilities in the methods section with “other type of feedback”.

We also completed the limitations section: “Finally, after reviewing the literature we decided to keep only five response options for the question assessing the reason(s) why GPs would find feedback useful, respectively seven options for the question about the type(s) of feedback they would like. These options seemed to be the most interesting, important or pertinent to study in the context of primary care. We decided to limit these response options, because with questions offering many possible response options, we were afraid that certain questions would have been completed incorrectly, for example at random. Though this restricted selection could result in a certain degree of information bias, we do not believe that this was in fact the case. Indeed, though “other reason” and “other type of feedback” were response options proposed to responders, they were never selected by GPs in our sample. In addition, no GP (n=7) suggested in the pretest phase that we add other response options to the list.”

DATA ANALYSIS:
1. Why was logistic regression used when a continuous Likert scale was available as the outcome measure? By dichotomizing the outcome variable to fit the data requirements of logistic regression, valuable information may have been lost.

We agree with the reviewer; valuable information may be lost by dichotomizing the outcome variable.

In fact, we were interested in comparing two groups of GPs, those who found feedback very useful (i.e. very useful or indispensable) and the others (i.e. those who found feedback rather useful, little useful or useless). All analyses carried out in our paper were laid out at the outset. We did not want to perform any post-hoc analyses, i.e. analyses which were not already planned in the study protocol. In particular, we were not really interested in assessing each sub-group of GPs according to the response to this question (feedback found indispensable, very useful, rather useful, little useful, useless).

Logistic regression was the type of multivariate approach that we planned in the study protocol. We believe that comparing those who found feedback useful and those who found feedback useless through logistic regression could help the reader to have a good grasp of the results. Table 4 shows (we think relatively clearly) non adjusted and adjusted ORs for the usefulness of feedback according to GPs’ socio-demographic characteristics. With five levels of response this multivariate analysis would have been more complex, and the results would have been more difficult to interpret and understand.

However, we decided to add in the results section the % for each sub-group of GPs according to the response given to the question about feedback usefulness: “11.6% found feedback indispensable, 32.7% very useful, 36.7% rather useful, 10.0% not very useful, 9.0% useless”

2. The use of stepwise regression seems unnecessary when sufficient literature exists to make educated theoretical hypotheses about the most likely predictors of usefulness perceptions, which would avoid the error inflation normally characteristic of automatic variable selection procedures like stepwise regression.

We planned to use stepwise regression, because, as stated in the introduction section, “several studies showed its [=feedback] effectiveness in preventive care, but it is currently not known whether GPs find feedback useful and, above all, what type of feedback intervention they actually would like to receive”. Though many authors have shown that feedback may be useful to improve health care, including in prevention, we did not know to the best of our knowledge if GPs found feedback useful. Therefore, it would have been difficult to make theoretical hypotheses about the most likely predictors of usefulness perceptions.

We completed the limitations section: “We used a stepwise selection procedure rather than a hypothesis driven selection for the multivariate analysis of usefulness perceptions, which could lead to a model giving an over-optimistic impression despite the additional fitting error this may have added. Though many authors have shown that feedback may be useful to improve health care, including in prevention, we did not know to the best of our knowledge if GPs found
feedback useful. Therefore, it would have been difficult to make theoretical hypotheses about the most likely predictors of usefulness perceptions to include in the model.”

3. With a 53% rate of non-response (which is not inherently a fatal flaw), I would have liked to see some manner of non-response analysis, even a cursory one.

Unfortunately, we cannot assess the differences between responders and non-responders, because we did not collect any data on non-responders. We added this limitation in the limitations section: “As we did not collect any data on non-responders, we could not assess potential differences between responders and non-responders”.

However, our sample appears to be relatively representative of all community-based GPs practising in Switzerland and France (data from Pays de la Loire). We completed the results section: “Our sample appears to be relatively similar in age and gender to all community-based GPs practising in Switzerland (professional organisation of Swiss physicians, 2016: median age 54 years (vs. 54 years in our study); men 59% (vs 61% in our study)) (ref) and France (Pays de Loire, 2013: median age 51 years (vs. 56 years in our study); men 57% (vs. 66% in our study) (ref)” and the limitations section: “However, our sample appears to be relatively representative in terms of age and gender of all community-based GPs practising in Switzerland and France (data from Pays de la Loire)”.

DISCUSSION/CONCLUSION:

1. The authors state on p. 13: "We found that younger GPs and those being more adherent to guidelines were more likely to consider a feedback useful. This suggests that they are probably more open to criticism and more receptive to feedback seen as a way of improving their practice." Though certainly a logically reasonable supposition, this conclusion needs support from the literature, as it is too big of a leap from what the results indicate.

We completed the discussion section: “Younger GPs (maybe because they are in general less experienced clinicians and, as a result, have not yet established their style of care) seem to have higher adherence to guidelines. (ref) For example, in a study assessing GPs’ attitudes to guidelines for elective surgical referral in England, it was shown that the odds of using guidelines decreased with increasing age, a ten year increase in age being associated with halving odds of use. (ref) By extension, it is hypothesized that younger GPs (and those being more adherent to guidelines) are also more receptive to feedback and therefore more likely to consider these interventions useful for improving their skills.”

2. Given the methodological details that need attention, I am not certain that the results and conclusion reached by the authors teach us anything particularly new or novel.

As this reviewer has previously stated, little is known on physicians’ acceptance of clinical performance feedback. (ref Payne & Hysong, 2016)
We stated in the introduction section: “several studies showed its [=feedback] effectiveness in preventive care, but it is currently not known whether GPs find feedback useful and, above all, what type of feedback intervention they actually would like to receive. It is very important to take GPs’ preferences into account in order to increase the acceptability of feedback interventions, and eventually increase guideline adherence in practice.”

We were pleased to identify that approximately half of GPs found that feedback would be very useful for evaluating their clinical practice for preventive care and that one of the two main reasons for feedback stated by GPs in our sample was to modify or improve their practice. These findings are encouraging, because they show that GPs are not reluctant to receive feedback interventions to improve their practice.

However, we also showed that they preferred to receive written (and not verbal) feedback: “very few GPs would like to discuss the results in local quality circle meetings, and even less face-to-face with the study investigators”.

Therefore, we think that our findings are important (they are also new, because, as stated in our conclusion, “the implementation of guidelines is highly related to the acceptance of feedback; we strongly encourage decision makers to take GPs’ preferences into account when developing strategies to implement guidelines, in order to improve quality in primary care”.

Minor changes:

1. p. 24, line 278 -- I presume the authors meant to say "odds ratio" rather than "odd ratio"

   Sorry for the typo!

2. There are various grammatical errors sprinkled throughout the manuscript; it may be useful to have the manuscript proofread and/or copy-edited by a colleague or a technical writer.

   Sorry for the grammatical errors! We asked a native English GP to check the paper.

Please also take a moment to check our website at http://imps.edmgr.com/l.asp?i=51746&l=60G1EX7Y for any additional comments that were saved as attachments. Please note that as Implementation Science has a policy of open peer review, you will be able to see the names of the reviewers.

If improvements to the English language within your manuscript have been requested, you should have your manuscript reviewed by someone who is fluent in English. If you would like professional help in revising this manuscript, you can use any reputable English language editing service. We can recommend our affiliates Nature Research Editing Service (http://bit.ly/NRES-HS) and American Journal Experts (http://bit.ly/AJE-HS) for help with English usage. Please note that use of an editing service is neither a requirement nor a guarantee of publication. Free assistance is available from our English language tutorial (https://www.springer.com/gb/authors-editors/authorandreviewertutorials/writinginenglish) and our Writing resources.
(http://www.biomedcentral.com/getpublished/writing-resources). These cover common mistakes that occur when writing in English.