Author's response to reviews

Title: Design choices made by target users for a pay-for-performance program in primary care: an action research approach

Authors:

Kirsten Kirschner (k.kirschner@iq.umcn.nl)
Jozé Braspennings (j.braspennings@iq.umcn.nl)
J.E. Annelies Jacobs (j.jacobs@iq.umcn.nl)
Richard Grol (r.grol@iq.umcn.nl)

Version: 2 Date: 4 February 2012

Author's response to reviews: see over
3 February, 2012

Dear Sir,

Thank you for giving us the opportunity to submit a revised version of our paper with the title “Design choices made by target users for a pay-for-performance program in primary care: an action research approach”. We would like to thank the reviewers for their comments which helped to improve the quality of the paper.

We made two documents:
- A file that includes additions and changes made in accordance with suggestions by the reviewers;
- A letter in which we describe in detail how we dealt with the comments.

We hope that our adjustments and comments are satisfactory and are looking forward to your reply.

Yours sincerely on behalf of the other authors,

K. Kirschner
Scientific Institute for Quality of Healthcare
Internal code 114
PO Box, 9101
6500 HB Nijmegen, The Netherlands
+ 31 24 3616338 (phone)
+ 31 24 3540166 (fax)
Referee comments and response

Reviewer 1
We appreciate the critical remarks of the reviewer, and will reply point-to-point below.

1. The reviewer indicates that more lessons should be learned from the QOF. Especially attention should have been given to the possible decline in non-incentivized aspects and to the (lack of) effect on intermediate outcomes. We discussed the design choices with the target users and brought in the available evidence. For instance, in the discussion on the payment it was made clear that research suggested that the payment should not be too high to avoid gaming, but high enough to make an effort for quality improvement. Although, the discussions took place before the referred papers were published, these subjects were addressed. GPs worried that the attention would be driven to the incentivized aspects of their performance, while P4P ideally had an effect on the overall quality of care in practice. The targets users therefore suggested to make a large set of indicators that could be incentivized in different baskets for different periods. This information is now described in the results “performance measurement” (section 1, line 7-12). We also discussed the possible effect of an increase on the process measures on the outcomes. GPs were convinced that the outcomes were a mix of patient and doctor’s performance. It was therefore decided that the payment should be based on the process measures indicating doctor’s performance, but the outcomes should be reported in the feedback as well. We added this statement to the results “performance measurement” (section 1, line 12-15).

2. The second point of the reviewer is about the remuneration. Questions were raised about the reimbursement formula, the seven levels of reimbursement, and the rewards based on improvement. We added formulas to the results “reimbursement” (section 1, line 5-15), a table on page 12 in which the bonus in the first year and the following years is presented for each of the seven levels and for the three domains. We also added an example of a bonus calculation for a mean practice with 2350 patient in a box on page 13.

The cost of the project was discussed by the target users. The payers maintained that the data collection should be paid for by the GPs themselves, and they paid a bonus based on the quality delivered. Since, the P4P standards were decided to be relative the amount of money to be allocated in the project was fixed; on average a practice with a 2350 patients could earn 7500 Euros (range: 0 to 15000 Euros). In the results on reimbursement we elaborated the text (page 11, line 34 and page 12, line 1).

3. The reviewer asks us how we suggest to deal with baseline performance that improves before introducing the program. If baseline levels are increasing in the same period as the P4P program then attribution is a huge problem in establishing its effect. As in the QOF our indicator set addresses chronic diseases such as diabetes, COPD and asthma, and cardiovascular risk management. Society is paying a lot of attention to these conditions due to the large numbers and cost resulting in a lot of projects to enhance its quality of care (effectiveness and cost). Knowing this we may expect that P4P is not the only initiative to improve quality of care on chronic disease during the study period. In our effect study (to be presented elsewhere) we will list other quality improvement projects of our participating practices to get more insight into this problem. The broader implementation of our P4P program will be based on the results of our effect study, but just as in the UK politics sometimes takes its own course.
4. A number of issues are raised under point 4.
   a. First of all, the reviewer suggests that participation is not an issue. “No practice would say no to a monetary incentive.” The reviewer questions the effectiveness of the bottom-up scheme, because the definition of effectiveness is not made clear enough. Bottom-up design choices means involving the stakeholders. Literature shows that this a necessary condition for sustainability (Gruen, 2008). We are not sure that money will have the same effect on sustainability. Furthermore, two recent reviews showed that bottom-up development of P4P increased its effect on the incentivized indicators. So, based on implementation science as well as the effect studies, we decided to develop a bottom-up program. We tried to make this more clear in the introduction (section 2, line 15-18).
   b. Another concern of the reviewer is the inclusion of process indicators and exclusion of outcome indicators in the P4P program. The GPs chose only process indicators for clinical care to be part of the P4P program, but they wanted to receive feedback on the outcome indicators as well. We made this more explicit in the paper on page 10 (section 1, line 12-15). This paper deals with the design of the program, but an effect study will show whether there is an effect on both process and outcome indicators.
   c. The reviewer sees a conflict of interest by involving the target users in the design choices. “The GPs could have chosen the easiest measures.” That is indeed an interesting question. A systematic consensus method (Delphi) was used to make the process of choice more transparent. This method is known in guideline and indicator development (Campbell, Braspenning et al, 2003) to avoid conflicts of interest. This line of reasoning has now been included in the introduction (section 3, line 26-29).

The reviewer stated that he is involved in research related to the Quality and Outcomes Framework in the United Kingdom. We agree with the reviewer that many lessons can be learned from this “experiment”. We are aware of the work provided by the University of Manchester, and work together with some of these researchers. Knowing that there is an urge for P4P, and learning from the QOF and the other scientific literature on pay-for-performance, we chose to make design choices together with the target users that is GPs and payers. We had a scientific interest in the process of making a bottom-up P4P program, and wondered about the effect of such a program. The paper presented here is on the development of the P4P program.
Reviewer 2
Reviewer 2 stated that this is a very interesting report of an innovative pay-for-performance program in primary care. The paper is organized well and written succinctly. No major concerns are addressed, but two points need our attention.

1. The reviewer suggests to insert Table 1 into the main manuscript. We inserted Table 1 on page 8 of the manuscript.
2. The reviewer would like to see some discussion about the implications of translation of our approach to other settings and geographic locations. We added this discussion to the discussion on page 15, line 4-12.