Author’s response to reviews

Title: Computerized clinical decision support system for diabetes in primary care does not improve quality of care: a cluster-randomized controlled trial

Authors:
Annemie Heselmans (Annemie.heselmans@gmail.com; annemie.heselmans@kuleuven.be)
Nicolas Delvaux (nicolas.delvaux@kuleuven.be)
Annouschka Laenen (annouschka.laenen@kuleuven.be)
Stijn Van de Velde (Stijn.vandevelde@fhi.no)
Dirk Ramaekers (Dirk.Ramaekers@kuleuven.be)
Ilkka Kunnamo (Ilkka.Kunnamo@duodecim.fi)
Bert Aertgeerts (Bert.Aertgeerts@kuleuven.be)

Version: 1 Date: 02 Oct 2019

Author’s response to reviews:

Dear editor,

We thank the peer-reviewers for their insightful comments! We hope that the changes we have made have further improved the manuscript. Below we provide a point by point answer to the peer reviewers suggestion. Changes in the text are indicated in red.

I confirm that all author details on the revised version are correct, that all authors have agreed to authorship and order of authorship for this manuscript and that all authors have the appropriate permissions and rights to the reported data.

Yours Sincerely,

Annemie Heselmans

Reviewer reports:

Reviewer #1: This is a manuscript that will be of wide interest to those interested in clinical decision support (CDS), and which reports the results of Belgian cluster-randomized trial of a
diabetes CDS intervention in primary care practices. I have several specific comments on the manuscript:

1. The CDS intervention did not improve either process or outcomes of care, principally because it was used in less than 2% of eligible cases. This point is not emphasized in the manuscript, and needs more emphasis including in the abstract/conclusions section, and the Discussion sections of the manuscript. It would also be helpful to note in the Introduction that careful perusal of prior CDS systems for outpatient diabetes care that failed (Meigs et al, Montori et al) had very low use rates, and systems that succeeded (K Peterson et al; O'Connor et al 2010 Ann Fam Med; Sperl-Hillen et al 2018 JAMIA; E. Kharbanda et al eGEMS) had high use rates. In the Discussion section I suggest mentioning that if a system is not used, you cannot tell if it works or not. First, you have to get it to be used.

Follow-up: We agree with you that a system that is not used in practice, could have an effect on improving care neither. We added this information and the relevant studies from the literature to the introduction (line 108) and made the manuscript more clear on this point.

Scripts were automatically triggered, and users could see that a script was available from their main screen, but have to click to view the script content. As such, click events on the tab of the decision-support could have given us an indication of physician’ interest in using the system. However, the coding of the registration of actual use was imperfect and did not take the click events into account, neither the number of times a specific reminder was triggered. The clicks were not registered and we had to base the analysis of actual use on the click events of different menu items or functionalities of the system. Because of this, the percentages did not reflect actual use, but use of the menu items of the decision support system. Unfortunately, because of this omission in programming of the number of clicks on the tab of the decision support, definite conclusions on actual use of the system could not be drawn based on the use of the menu items of the system. We do not know how many times reminders were used or read which is an important limitation of the study. (line 430).

We changed the wording from ‘functionalities’ to ‘menu items’ of the decision support systems through the manuscript and added examples of the menu items (eg disabling the reminder, giving information about the utility of the reminder, etc). We adapted the Methods section of the manuscript to make that point more clearly as follows: (starting from line 208):

A new record was inserted in the log file foreach different decision support (DS) message that was triggered. If the identification (id) code of the DS message already existed in the database, no new record was inserted. The group of DS messages can be classified in four categories: drug contraindications and interactions, links to guidelines and reminders. DS messages were automatically triggered, and users could see that a message was available from their main screen but have to click to view the script content. We did not register the number of clicks on the tab of the decision support (see Discussion section), but registered the number of clicks on the menu items or functionalities of the EBMeDS system. Following events were monitored: disabling a reminder, giving feedback about a reminder, giving information about the utility of a reminders, requesting details or information about the reminders.
The results section was changed as follows (starting from line 282): 322 and 354 different DS messages displayed per practice and per month, respectively during the first six months and between the sixth and the twelfth months of the study. During the first half year, 9.69% of all different DS messages were reminders, of which 5.66% were diabetes reminders. During the second half year, 11.18% of all DS messages were reminders, of which 5.63% were diabetes reminders. Menu items or functionalities of the decision support system (all clinical areas) were 127 times used during the first six months (or in at least 1.3% of cases) and 120 times between the sixth and the twelfth month (or in at least 1.2% of cases). The utility of the diabetes reminders was given 4 times during the first six months. The other menu items were not used within the group of diabetes reminders during the follow-up period of the study.

Because no definite conclusions can be drawn based on the data of actual use, the issue of low use rate was deleted in the abstract line 52.

2. Please comment in the Discussion section on why the CDS was infrequently used. There are clues in the survey data that are included in this manuscript. Focus sharply on this question, especially if you want to do a follow-up study with a system that may have a better chance of working.

We added the following text in the Discussion section (line 441):

Based on the interviews we had with the end-users, we could assume that there was a lack of use of the system because of the following reasons:

- The need to invest time in the system;

- The technical problems that occurred in larger practices because of the need to activate the datawarehouse to allow automatical export of data;

- The shortcomings in the organizational context of Belgian primary care for chronic disease management;

- The lack of relevance of the reminders mainly because of coding problems, could be possible reasons for the lack of use of the system.

We added the following wording (line 368): Based on the systematic review of Van de Velde et al, a 16-factor checklist was developed consisting of four domains which may affect the use of the system and success of the effect of CCDS systems (context, content, system and implementation).

3. A prior publication in Implementation Science provides more detail on this project. However, it is very important that the present manuscript include a paragraph (and maybe some Figures) that describe and illustrate the actual content of the CDS as seen by the primary care clinician. The absence of this in the present manuscript sends the reader on a wild goose chase through the literature. Please fix this major omission.
Follow-up: The content of the diabetes reminders and screenshots of the EBMeds system are displayed in additional file 3.

4. The 6-month and 12-month follow-up periods are too short to detect changes in glucose control and lipid control. This is a design limitation of the study. Please provide the reader with the average length of time between the initial and follow-up measures for BP, A1c, and LDL. These data are very important. Since the study was only 12 months long, and the patient's first visit may not occur at the go-live date for the CDS, the mean follow-up time is going to be too short to detect changes (plus the 2% use rate). Please also indicate whether (a) the analysis was limited to those that had a second visit, or (b) what you did with those who only had one visit, or one A1c or LDL value—state whether you carried the last value forward.

Follow-up: All patients with diabetes type II were included in the study regardless of their number of visits during the follow-up period. We are aware that there might be patients with no second visit, but we assume that these patients were equally distributed between the intervention and control group because of the randomization. Pre- to post implementation changes in HbA1c and LDL were calculated as the difference between the mean value of HbA1c and LDL of the previous year at the moment of the start of the study and the mean value of HbA1c and LDL of the previous year at the moment of the study end. As such, mean follow-up period for each patient was one year. This information was added in the Methods and the Discussion section (line 144, 149, 19, and line 477).

The limitations of the short follow-up period to detect significant changes in glucose and lipid control was added in the Discussion section (line 476).

5. The manuscript is unclear about the "intent-to-treat" analysis. The abstract says 51 PCPs were randomized, but the results appear to include data from only 29 practices. Line 394 and Line 381 indicate why some practices were lost. Please consolidate this information in one place.

Follow-up: We agree with this point of confusion.

The information is now consolidated in the Discussion section as follows (line 450–470) and mentioned in the Results section (line 244):

We were ready to start in the beginning of 2014 but had to postpone the study start due to technical problems and software difficulties with exporting the data from the EHRs. Software problems were not immediately solved because of other priorities and accreditation deadlines within the software firm. Communication of the problems with the trusted third party was not easy and delayed the project considerably. During the time of delay, 3 practices (12 physicians) changed their software of EHR and could not participate anymore.

We noticed a considerable level of drop-out (43%). Drop-out rate of practices were similar between the IG (44%) and the CG (42%), but relatively larger practices in the intervention group
dropped out, resulting in a difference of 10 physicians and 1113 patients between the two groups, with more physicians and patients in the control group, (21 vs 31 physicians and 1351 vs 2464 patients). The system requirements significantly slowed down the EHR of larger practices which was an important reason to deactivate the EBMeDS system. Although analysis was done following the intention-to-treat principle, patients in these practices could not be analyzed anymore because these practices stopped their participation in the trial.

6. It would be a great contribution to the literature to create a "box" that lists major challenges that eroded the success of the intervention such as: follow-up time was too short; practices drop out because of IT problems; very low use rates; some docs felt CDS was inaccurate (was it inaccurate or incomplete?); etc.

Follow-up: A summary of possible reasons for failure of the system can be found in the Discussion section (line 485).

7. Only about 100 diabetes patients per practice were analyzed. Were there any exclusion criteria for using a patient in the analysis? Please state them (age, type of diabetes, number of visits, etc.).

Follow-up: All patients with diabetes type II and 18 years or older were included in the trial. Inclusion criteria were mentioned in the Methods section (line 144). There were no limitations on the number of visits. This was also added to the Discussion section (line 477).

8. You report changes in A1c, LDL, and BP for all patients. However, at baseline many patients were already at goal. Please add additional results reporting the changes in A1c, BP, and LDL only for the subset of patients who were above clinical goal at baseline.

Follow-up: Post-hoc analyses of the subgroup of patients with a mean HbA1c $\geq$ 7% at baseline revealed a statistically significant difference in pre- to postimplementation change of HbA1c of -0.40 (95% C.I. -0.70; -0.09) or an effect size of 0.31 in favor of the intervention group after twelve months of follow-up. Post-hoc analyses of the subgroup of patients with LDL $\geq$ 100 mg/dL and blood pressure $\geq$ 130/80 mm Hg did not show statistically significant differences in pre- to postimplementation change between groups in LDL cholesterol and blood pressure. See table 5.

This information was added to the Results section (line 275). The post-hoc analysis was added to the Discussion section (line 361).

9. Line 137: Please state how you defined the "baseline time point."

Follow-up: Baseline time point was defined as the date of the first export. This information was added (line 149).
10. Line 143: The CDS included "diabetes specific reminders and suggestions." Please expand on this and show the interface the clinicians saw. State whether the CDS was designed to be given to or shared with patients, and whether the CDS suggestions were prioritized in any way, or listed in a certain sequence. Generally speaking, prioritized CDS is more useful than non-prioritized CDS to primary care clinicians.

Follow-up: The decision support is only targeted at the physician and was not designed to be shared with patients.

The EBMeDS system covers a broad spectrum of clinical areas. Reminders are prioritized. Urgent reminders are displayed as pop-ups in red (e.g. high potassium values when taking spironolactone). Reminders that focus on clinical outcomes (such as e.g. adding ACE-inhibition in diabetic nephropathy, prescribing aspirin in vascular disease, etc) are indicated in yellow and are placed at the top of the list of reminders. Reminders that are focused on process or surrogate outcomes (e.g. LDL concentration, HBA1c control, etc) are shown in gray and at the bottom of the list (line 165). A screenshot of the interface and a complete list of reminders related to diabetes are displayed in additional file 1.

11. In addition to diabetes CDS, how many other CDS suggestions were provide at each visit, on average. You give (line 243 ff. the mean number of diabetes CDS messages per practice per month. How many were delivered at the average diabetes visit of a single patient? In my medical group, reminders have been deemed ineffective except for cancer prevention and immunizations, and we are actively eliminating reminders.

Follow-up: We agree that this is an important question. However, it was not possible to extract the number of messages at an average diabetes visit from our database.

As also mentioned in the follow-up of comment 1, we encountered different problems because of shortcomings in the code of registration of actual use. We have rewritten the Methods, the Results and the Discussion section to make this point more clear (see follow-up on comment 1)

Before the start of the study, three general practitioners selected relevant reminders for general practice independently from each others. This information was included in the Methods section (Trial preparation line 157).

12. Line 3245: Many clinicians deemed the reminders to be "irrelevant." Say which ones were most irrelevant. If you were revising the CDS, how would you re-design it? That advice could help readers a LOT.

Follow-up: Discussion about the issue of relevance of reminders learnt us that relevance was not related to specific reminders, but to coding problems in general. E.g. reminders showed up to test HbA1c while it was already recently done but not correctly captured by the system because of incorrect coding. This issue was made more clear in the Discussion section (line 375).
We assume that the lack of improvement of the system was mainly caused by imperfections in the organizational context of Belgian primary care for chronic disease management and shortcomings in the system requirements for the correct use of the EBMeDS system (such as e.g. complete structured records), rather than in the design of the EBMeDS system, (line 501)

Only one physician was not satisfied with the EBMeDS system. Negative comments about the intervention during interviews were seldomly made on the level of the design of the EBMeDS system which makes us presume that future intervention priorities in Belgium must focus on the creation of the right conditions for successful implementation of computerized decision support systems in the first place (line 505).

13. Did you do a pilot test? Why not?

Follow-up: We did a pilot test in another Belgian EHR. This was mentioned in the Introduction (line 120): ‘Based on the results of a previous qualitative study in a pilot setting, we adapted the system conform the reflections of the end-users which is an important requirement for the acceptability of the system.’

See also in the Discussion section (line 421).

14. I hope you will re-load and try again.

15. You need some newer references.

Follow-up: Following references were added:


Sperl-Hillen JM, Rossom RC, Kharbanda EO, Gold R GE, Elliott TE, Desai JR, et al. Priorities Wizard: Multisite Web-Based Primary Care Clinical Decision Support Improved Chronic Care

Reviewer #2: This is an article to which I do not have much to add. Punctuation and description can be accepted in this format. In my view, it is very clear, the definition of the Sample, the significance and the Power test. I only request that the Effect Size be reported.

Follow-up: The effect size on which the sample size was calculated was reported in the Methods section (line 217). Effect sizes of the effect of the intervention on the primary outcome were reported in the Results section (line 258 and line 277). Effect sizes of the effect on secondary outcomes were not reported because no statistically significant effect was found on these outcomes.