Author's response to reviews

Title: The association between improved quality diabetes indicators, health outcomes and costs: towards constructing a "business case" for quality of diabetes care- a time series study.

Authors:

Rachel Wilf-Miron (r.w.miron@gmail.com)
Arkady Bolutin (arkadyv@bgu.ac.il)
Nesia Gordon (gordon_ne@mac.org.il)
Avi Porath (porat@bgumail.bgu.ac.il)
Ronit Peled (ronitpeled@yahoo.com)

Version:2 Date:14 August 2014

Author's response to reviews:

August 10 ,2014

Editor in chief
BMC Endocrine Disorders

Re: Our article The association between improved quality diabetes indicators, health outcomes and costs: towards constructing a "business case" for quality of diabetes care- a time series study.

Dear Editor,

We are pleased to re-submit our manuscript The association between improved quality diabetes indicators, health outcomes and costs: towards constructing a "business case" for quality of diabetes care- a time series study.

to your journal after addressing all the reviewer's comments. All changes in the manuscript are marked in yellow as a respond to referee no. 1 and in green to referee no. 3. Referee no. 2 had no comments. In gray we have marked the sentences that were added according to the editor's special request.

Hereby are the referees' comments:

Referee #1:
Major compulsory revisions

1) I would say HbA1c>9% is very poor rather than poor. The thresholds used needs to be put into context: US, UK (NICE) or WHO recommendations? A reference for the definition has been added; however, MHS as well as the Israeli National Quality Scheme (Manor O, Shmueli A, Ben-Yehuda et al. Quality Indicators for Community Care in Israel. Public report (2008-2010). Available at: http://www.health.gov.il/Publicationsfiles/qindicatorsReport2008-2010.pdf )
define HbA1C >9% as "poor" and not "very poor". This definition is also in accordance with the American National Committee for Quality Assurance (HEDIS) definition which says in "The state of health care quality 2012, pg 57, under measure definition: "HbA1c poor control (>9%)".

The British Quality and Outcome Framework also uses the threshold of (>9%) as a measure of diabetes care (see for example measure DM25 in the 2009-2010 report, available at http://www.hscic.gov.uk/catalogue/PUB04431/qof-09-10-rep.pdf.

In addition, the poor diabetes control (HbA1C >9%), was selected as a quality indicator for diabetes care, by the National Diabetes Quality Improvement Alliance, to be used among the the OECD countries. Please see at: (http://intqhc.oxfordjournals.org/content/18/suppl_1/26/T1.expansion.html)

I would be more interested in the risks associated with very low HbA1c levels as well since these are associated with hypoglycaemia. In general there appear to exist U-shaped relationships between these biological parameters and outcomes (hospitalisation, survival) and you seem to be the very high values vs all others. It would be interested to see categories rather than a dichotomised variable. How did the authors decide on the thresholds they used? We would like to add to the above explanation that we have been aware of the new evidence that emerged around 2007, for example in the publication of Quassem et al. (Ann Internal Medicine, 2007;147:417-422). These publications led to the redefinition of HbA1C control, which until 2009 (including for example HEDIS and National Israeli reports from 2009 ) was defined as < 7%, and from 2010 is defined as follows: for patients aged 0-74 adequate control is <7% and for patients aged 75+ adequate control is < 8%. This shift in definition for the 75+ year old patients resulted from balancing the risks of hypoglycemia and it's consequences with the long term benefits of strict control (<7%). As stated in the article, we chose for the analysis only measures that were "stable" i.e. its definition did not change during the study period. Adequate control, which was redefined during those years, was therefore considered less suitable for our analysis.

2) The authors give a decent overview of the informatics revolution but fail to mention the Quality Outcomes Framework, the largest incentivisation and performance monitoring system in the world, while submitting a paper on quality of care to a British journal. Incentivisation and monitoring though the system has led to improvements in quality of care (http://www.ncbi.nlm.nih.gov/pubmed/19625717, http://www.ncbi.nlm.nih.gov/pubmed/21712336), reduced inequalities (http://www.ncbi.nlm.nih.gov/pubmed/18701159) and these effects were also observed for diabetes patients (http://www.ncbi.nlm.nih.gov/pubmed/22918988) while hospitalisation for diabetes patients who received better QOF care appeared to be lower (http://www.ncbi.nlm.nih.gov/pubmed/20880046). Finally, the measurement indicators have more or less been removed from the scheme, when a 'paired' control indicator is still in, and it has been shown that the levels of
care were not affected by the withdrawal of indicators (http://www.ncbi.nlm.nih.gov/pubmed/24468469) and the control indicators should be the focus. Also note that the diabetes clinical domain of the scheme was the biggest with 17 clinical indicators (now reduced to 15 I think but still the biggest). All this is relevant (especially the last three papers) and I would expect the authors to dedicate a para in the introduction to discuss the QOF and its diabetes domain (and perhaps overall as well). Done – two of the above mentioned references have been added to the background.

3) The section regarding GEE in the statistical analysis section is unclear: what is the purpose of the correlation structure i.e. where is it used, what does ‘within-group’ refer to i.e. what are the groups? If this is relevant to the nested structure of the dataset, which has not been properly explained, see point 4 below. The role of the global and dynamic models and why their coefficients difference answers the question is unclear and needs to be properly explained. Done - the section describing the GEE modeling has been completely revised.

4) The nested structure of the data needs to be properly explained. Lines of data are not an elaborate way to do so. For example I would say “The data had a nested structure with patients nested within areas (if relevant) and crossed with year (since same patients more or less over different time points, rather than different). Also the Poisson and Cox regression analyses need to account for the structure of the data and robust standard errors need to be calculated. In Stata clusters for the patients need to be defined with xtpoisson and the vce(cluster) option (the latter will allow robust error calculation within patient, who have been defined in stset using the id option). Of course all this relates to Stata and I don’t know what the authors used for their analyses. Done – please see the revised section on statistical methods.

5) Discussion is too long a bit all over the place. Please use a standard reporting: such as : key findings / strengths and limitation of the study / results in context (i.e. in relation to other work) / conclusion. And try to reduce a bit. Done

6) Results section is poor. The results are not presented clearly enough and the section needs to be expanded. For example there is no reference to the results from the Poisson regression (table 3) other than mention of a significant association. In sharp contrast to the massive discussion section. Also report incidence-rate ratios rather that betas for the Poisson analyses, since the irrs can be interpreted much more easily. Done

Minor essential revisions. 1) Abstract introduction is unclear. Grammatically, needs “constructs” instead of “construct” but still doesn’t make much sense to me and I think needs to be rephrased. i.e. replace “business case” which would need to be defined with something like “cost-effectiveness”? Done

2) List covariates in the methods section of the abstract. Done

3) Mention and reference the software used. Done
4) Interaction term is first mentioned in the results, not the methods – relates to main point 3. Done

5) In the results section report confidence intervals rather than standard errors. Done

Referee No. 2: had no comments

Referee No. 3:

Overall impression of the study:
The objective of the study is clearly defined. However, the discussion and conclusions drawn from the results as well as title and abstract do not fit well to the objective.

Declared objectives of the study are:
1) investigation of the relationship between improvement in selected measures of diabetes (type 2) care and patients' health outcomes
2) to estimate the effect of improved performance on direct medical cost.

Providing the base for a “business case” is not a declared objective of the paper and the introduction does not allude to any potential business case.

Moreover, the results are too insufficient to be used for serious business case development. Business case development in health care is very complex, but the present study provides nothing but a single estimate of potential cost savings. It is an exaggeration to state that the "result (...) is a preliminary pivotal for proving such a business case" (Discussion section, paragraph 12).

Nevertheless, title and abstract of the paper take up this issue and suggest that a business case has been developed and that important conclusions for international health care systems other than the MHS system could be derived.

The latter suggestion relies on a single phrase of the conclusion-section of the paper, saying: “In health care systems with relatively low member turnover, this finding presents an important milestone linking quality and cost, helping to construct the business case approach to quality”. It is not defined in any way what a “health care systems with relatively low member turnover” is and why the results of the present study should be representative of these systems and not of others. An explanation has been added.

The conclusion section and the abstract suggest that the findings of the study are generalizable but this conclusion is without any foundation.

In sum, the conclusions are not adequately supported by the data. The conclusion has been revised.

The data themselves are another issue: Superficial, fragmentary and misleading descriptions of the methods and results make evaluations of their soundness and appropriateness of measures difficult.

Even though the “Methods” section is well structured (Setting, Focus of analysis, Study population, Data source, Quality indicator Definitions, Statistical Analysis), methods are not well described and many questions arise: The statistical
methods section has been revised.

• The calculation of cost savings was obviously based on a price list published by the Israeli Ministry of Health. The price list is neither referenced nor described in detail. What kinds of cost (direct cost, indirect cost…) were actually included in order to determine the costs of a hospitalization day? How much was a hospitalization day assumed to be? Did prices change during the observation period or were they assumed to be stable? A reference to the price list and an explanation have been added.

• What led to the expectation that improved control in 2009 had an effect on hospitalization at least until 2012?

There is no "magic formula" to be used to calculate how many years may pass between the investment in improving quality of care and an observed improvement in outcomes. We do know that the process indicators improve immediately; however, complications and undesired outcomes may be observed within a few months or years. The literature suggests mostly a gap of 3-5 years until the investment in quality can be harvested (T.P. Gilmer et al., “Impact of Office Systems and Improvement Strategies on Costs of Care for Adults with Diabetes,” Diabetes Care 29, no. 6 (2006):1242–1248). This is the main reason why we have decided to prolong the study period until 2012 (this was the year when the data set was constructed).

• In order to estimate the effect on death “data on death event from the years 2010-2012 were added” (Statistical analysis, paragraph 3). Where were these data taken from? Explained

• Some statistical tests are obviously used, but not described in the “Methods” section. For instance, Table 2 presents results of “linear regression models for curve fit”, but such models are not described in the “Methods” section. Done

• Why a Poisson regression model is used in addition to GEE models? Explained

• What is meant with “Global and Dynamic GEE models”? Explained

• The description of statistical methods is not consistent with the description of the results. For instance, the “Methods” section enumerates three different statistical methods that were used: Poisson regression, global GEE models and dynamic GEE models. However, the text of the “Results” section describes no results of the Poisson regression. Only table 3 presents results of the Poisson regression, but the text referring to table 3 describes results of the global GEE model. Done

Concerning the description of results, further questions arise:

• According to the methods section, patients who were diagnosed during the study period were included. This raises the question concerning the mean period of observation of patients. This important information about the study population's baseline characteristics is lacking.

Because each subject entered the study in a different time and stayed for a different period of time, a time series approach was used to address this unique panel structure. Please see statistical methods revised section.
• A description of the outcome variables’ distribution is lacking. How many hospitalization days/ED visits were counted per year in total and in average per patient? How were outliers handled? How many death events were counted?

• Models for HBA1c testing and ED visits did not reach statistical significance, but the respective results of statistical analysis are not displayed. Done

Some limitations of the study are stated. However, linguistic problems impair comprehensibility. Done

• “(3) One can argue that improved health outcomes and reduced cost result from other factors, such as diabetes co-morbidities, such as cardiovascular disease and personal characteristics and life-style. MHS internal data shows that from 2003 to 2009, the proportion of adult members with any cardiovascular disease has been stable.” (Discussion section, paragraph 15)

It is not clear how diabetes co-morbidities should improve outcomes. Maybe the authors mean “the treatment of co-morbidities”? The proportion of adults with any cardiovascular disease is not interesting, given the fact that cardiovascular disease is a chronic condition. Provided that the number of patients in the MHS Diabetes registry has been more or less stable (actually we cannot know that, because the authors don’t provide any information about that), the number of patients with cardiovascular co-morbidities will always be stable, too. What’s interesting, is the question whether patients are (successfully) treated for their co-morbidities or not?

We have decided to present our internal information, which indicates that the cardiovascular registry (as a marker for co-morbidity) had been stable for the study period. In a way, this argument controls or adjusts our panel for co-morbidity.

• “(4) with the lack of a control group, the effect of the improvement in the quality indicators on outcomes and cost could be measured.” (Discussion section paragraph 15) If it could be validly measured without control group, why should someone demand a control group? Or could it NOT be measured? If it could not be measured, all the results of the study must be questioned. Explained

Major Compulsory Revisions:

1) Resolve inconsistencies between objectives, discussion and conclusions; either drop the “Business Case” subject and reframe discussion and conclusions Done

or point out:
- the clear definition of the terms “Business Case” and “Health system with low enrollee turnover” Done
- which stakeholders of a health system would benefit from such a “Business Case” - what else besides the results of the study, would be necessary to develop such a “Business Case” Done
- why and under which circumstances the results of the study are generalizable Done
2) Adapt abstract and title according to 1)
3) Reference and describe the price list of the Israeli Ministry of Health in detail. Done

Point out:
- What kinds of cost (direct cost, indirect cost…) were included in order to determine the costs of a hospitalization day? Done
- How much was a hospitalization day?
- Did prices change during the observation period?
4) Point out what led to the expectation that improved HbA1c control in 2009 had an effect on hospitalization at least until 2012 Explained above
5) Where were data on death events between 2010 and 2012 taken from? Explained
6) Add a complete description of all statistical analyses used in this study to the Method section. Done
7) Point out why Poisson regression model was used in addition to GEE models. Done
8) Define what is meant with “Global and Dynamic GEE models” and why they were used. Done
9) Resolve inconsistencies between text (Results section paragraph 3) and contents of Table 3. Done
10) Add a description of the period of observation per patient to the description of the study population’s baseline characteristics (at least mean/standard deviation) Done
11) Describe the distribution of outcome variables, at least:
- How many hospitalization days/ED visits were counted per year in total and in average per patient? - How were outliers handled?
- How many death events were counted? Explained in the statistical methods in details.
12) Display results of statistical analyses concerning HbA1c testing and ED visits. Done
13) State more precisely what is meant by “One can argue that improved health outcomes and reduced cost result from other factors, such as diabetes co-morbidities, such as cardiovascular disease and personal characteristics and life-style. MHS internal data shows that from 2003 to 2009, the proportion of adult members with any cardiovascular disease has been stable.” (Discussion section, paragraph 15) Done
14) State more precisely what is meant by “with the lack of a control group, the effect of the improvement in the quality indicators on outcomes and cost could be measured.” (Discussion section, paragraph 15) Done

Minor Essential Revisions:
1) Table 3): Add explanation of the abbreviation S.E. (standard error) Done
2) Table 6): Add explanation of the abbreviation NIS (New Israeli Shequel) Done
3) Check orthography: “The quality of health care has drawn increasing attention form health care systems during the last two decades” (Background section, paragraph 1) Corrected
4) Check grammar: “In 2004, MHS implemented a “Performance Measurement System” (PMS) that help focus the organization’s attention on selected clinical domains including diabetes care.” (Background section, paragraph 3) Done
5) Decide whether to use “data” as singular or plural word and use it consistently: “For this task data on death event from the years 2010-2012 were added.” (Statistical Analysis section, paragraph 4); “Although data on hospitalizations specifically related to diabetes are not available, crude national figures indicate a steady increase in hospitalization days between 2000 and 2010 (22).” (Discussion section, paragraph 6); “Data from the United Kingdom has shown that the mortality risk among patients with type 2 diabetes is 1.6 times higher than that of the general population (24).” (Discussion section, paragraph 9); “MHS internal data shows that from 2003 to 2009, the proportion of adult members with any cardiovascular disease has been stable.” (Discussion section, paragraph 15); Done
6) Avoid using abbreviations of verbs (can’t instead of cannot): Done
“hence, the data can’t be appropriately used to substantiate returns on investments in quality.” (Discussion section, paragraph 15)

We would like to thank the reviewers for their comments which helped us to improve our manuscript and for your consideration.

Looking forward to publishing in your journal,

Sincerely,

Dr. Ronit Peled
Health Systems Management,
Ben Gurion University of the Negev.