Author's response to reviews

Title: Q-Score: Development of a new metric for continuous glucose monitoring that enables stratification of antihyperglycaemic therapies

Authors:

Petra Augstein (petra.augstein@gmx.net)
Peter Heinke (heinke@diabetes-karlsburg.de)
Lutz Vogt (vogt@diabetes-service-center.de)
Roberto Vogt (rvogt@ibvnet.de)
Christine Rackow (chrackow@web.de)
Klaus-Dieter Kohnert (kohnert@diabetes-karlsburg.de)
Eckhardt Salzsieder (salzsied@diabetes-karlsburg.de)

Version: 4 Date: 13 March 2015

Author's response to reviews: see over
Dear Editor,

Thank you very much for the review of our manuscript, “Q-Score: Development of a new metric for continuous glucose monitoring that enables stratification of antihyperglycaemic therapies”. We have revised our manuscript in the light of the reviewers’ comments and made the required changes. Please find below our point-by-point responses to the reviewers comments, with the locations (line numbers) of new text highlighted in yellow.

We greatly appreciate your support and look forward to hear from you,

Sincerely,

Petra Augstein

Point-by-point responses.

Reviewer 1

Point 1
They do not adequately acknowledge the existence and/or limitations of the work of other authors who have utilized other patter-recognition approaches to classifying CGM data with the intent of providing personalized care tools.

We appreciate recognition of our work as an “inventive application of factor analysis to a meaningful clinical issue” and the suggestion to acknowledge related studies. We performed a new literature review. In the revised manuscript we discussed the work of Rawlings et al. and of Marling et al. [1, 2] (lines 80–90). The first group developed a graphical user interface applying glucose variability metrics for the assessment of CGM profiles and tested it in a proof of concept study. The second group developed a consensus perceived glycemic variability metric that captures the gestalt perceptions of 12 experienced physicians actively managing patients with T1DM. We also included the work of Fabris et al. [3] (lines 292–309), who performed a Sparse Principal Component Analysis of parameters of glucose variability. The suggested reports by Eberle et al. and Jensen et al. described the development of models of insulin and glucose metabolism intended for the prediction of glucose values, the comparison of prognosis with actual outcomes and the application of the models. The Q-Score allows the assessment of the quality of the actual glucose profiles and is not a model.

Point 2
Central to the difficulty in evaluating the validity of the present approach is the lack of information about the dataset used for validation and the lack of traditional metrics for the accuracy of classification tools. It is not clear whether the evaluation of the performance of the Q-score was conducted using CGM data from the same set of records that were used to
develop the score or whether those evaluation records were from an independent sample. As such, it is not possible for the reader to make a judgment about the appropriateness and extent of validation of the Q-score.

All 1,562 CGM profiles were used for the development of the Q-Score. A subset of all CGM profiles (n = 766) was assessed independently by three diabetes specialists, who categorised the blood glucose profiles based on their own clinical experience. The criteria for the categorisation are given in Figure 1B. However, as the evaluation of the CGM profiles was subjective, a different categorisation of CGM profiles is possible depending on the assessment of the physician. We thank the referee for the comments on the validation of the Q-Score and have revised the manuscript accordingly (lines 133–135).

Point 3

Once it has been clarified whether the Q-score was validated on new records, or on a subset of the records used in score development, the method of validation of the score should be formalized. Some ‘flavor’ of sensitivity and specificity (correct and false categorizations) should be used. We refer the authors to the two citations above and to Wang Y, Wu X, Mo X., A novel adaptive-weighted-average framework for blood glucose prediction. Diabetes Technol Ther. 2013 Oct;15(10):792-801. doi: 10.1089/dia.2013.0104. Epub 2013 Jul 24., for examples of evaluation metrics for classification.

The validation of the Q-Scores including calculation of sensitivity and specificity requires a suitable method or procedure. Our only comparison or validation group available are the 766 CGM profiles categorised by the three diabetes specialists. The categorisation is subjective and can differ according to the clinical and practical experience of the clinician. Therefore, we assessed the inter-rater reliability using Cohen’s kappa score [4] to determine the concordance of assessments of CGM profiles by Q-Score and by diabetes specialists. Cohen’s kappa score revealed substantial concordance between Q-Score and the physician’s assessment for two diabetes specialists using the scale of Landis and Koch [5] (Physician A: 0.759 ± 0.015; Physician B: 0.724 ± 0.015) and moderate concordance for the third diabetes specialist (Physician C: 0.519 ± 0.018). This indicates complete concordance for the categorisation of 59.1% of the CGM profiles, a deviation in categorisation by one level in 37.4% of CGM profiles and by two levels in 3.5% of CGM profiles. These data are included in the revised manuscript as a new paragraph (lines 192–213) and the methods have been adapted (lines 144–147).

Point 4

As presented, the Q-score would be limited in its clinical utility as the reader lacks information about whether this scoring method could be applied to any CGM and what the magnitude or nature of possible errors related to the application of the score to categorize glucose control for a single new individual by a physician in practice.

The Q-Score can be applied to any CGM. The clinical utility needs to be tested in a future study.
Point 5

Why were the individual components of the Q-score standardised? The factors that they represent accounted for different portions of the total variance in CGM profile...were weighted or other modifications of the Q-score parameter considered and compared to the Q-score with standardised components?

From all factors, we selected one parameter that had a high factor loading, was simple to calculate, and was easy to interpret in the context of a CGM curve for general practitioners. These parameters were the MBG and the time spent above the target range from factor 1; the range from factor 2; the time spent below the target range from factor 3; and the MODD from factor 4 (Supplementary Figure 2).

The five selected parameters have different distributions with unequal means and variances as well as different units. To achieve equivalence in the parameters for calculations, the five selected parameters with unequal means and variances were standardised with a z-transformation. The Q-Score was computed as the sum of all five standardised variables. This ensured that all five parameters had an equal impact on the Q-Score. Then, to ensure positive values, we added a constant equal to 8.

Minor Essential Revisions:
1) The phrase “with principal component analysis” appears in the abstract, and “PCA, principal component analysis” appears in the list of abbreviations but nowhere in the manuscript. Principal component analysis is subtly different than factor analysis, and certainly not a part of factor analysis, so these references to it should probably be removed.

We agree with the reviewer and have omitted the phrase in the revised manuscript.

Reviewer 2

Point 1

The Q score from the CGMS data is correlated to patients metabolic control status. Metabolic control status was classified to very good, good, satisfactory, fair and poor by three diabetes specialists. But, there is no any definition about the classification. It could be very obscure. I think more accurate and concrete definitions for classification of patients according to metabolic control or glucose control should be shown in this paper.

Diabetes specialists assessed the CGM profiles based on their personal experience and opinions. The categorisations are explained in Figure 1B. We acknowledge the comment regarding the classification and added a reference to supplementary Figure 3.

We agree with the reviewer and have presented a table providing patient demographics and characteristics including BMI, HbA1c, daily insulin dose (if applicable) for type 1 and type 2 diabetic subjects in the revised manuscript. The type 2 diabetic subjects are shown as subsets according treatment (diet, OHA OHA with insulin, insulin alone).
Point 2

Mean HbA1c of the subjects analysed in the paper was 7.0%, which means that their glucose control status was relatively good. I think CGMS data are so variable according to glucose control status. For example, a patient with good glucose control status would show a small range of glucose change, while some patients with bad glucose control status would show a large range of glucose change. So the proportion according to glucose control status could be an important factor. I recommend Q score and its implication should be shown according to level of HbA1c. For example, HbA1c < 7%, 7~8%, 8~10%, >10%.

We appreciate this comment and could not agree more. Although the mean HbA1c was indeed 7%, the advantage of our study is the large number of investigated CGM profiles. Therefore, all Q-Score categories were representative of a large number of cases (very good to fair >350 per category, poor >170 cases).

The reviewer is absolutely right; the Q-Score is related to the glucose control status, and increases with HbA1c (please see the table below). Nevertheless, the minima and maxima of the Q-Score in the HbA1c classes reveal that very good as well as very bad Q-Scores are present in all HbA1c classes (please see table below).

Table: Q-Score in relation to HbA1c classes

<table>
<thead>
<tr>
<th>HbA1c (%)</th>
<th>n</th>
<th>Q-Score Mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 6.5</td>
<td>531</td>
<td>6.23 ± 2.77</td>
</tr>
<tr>
<td>6.5 – 7.0</td>
<td>375</td>
<td>7.56 ± 2.93</td>
</tr>
<tr>
<td>7.0 – 7.5</td>
<td>322</td>
<td>8.62 ± 3.17</td>
</tr>
<tr>
<td>7.5 – 8.0</td>
<td>155</td>
<td>9.96 ± 3.33</td>
</tr>
<tr>
<td>&gt; 8.0</td>
<td>179</td>
<td>11.72 ± 3.68</td>
</tr>
</tbody>
</table>

We have revised the manuscript (lines 229–234) and added a new table (Table 3) in accordance with the advice of the reviewer.

Point 3

Use of Insulin could affect glucose change so much. I can not see medication history of the patients. Medications of the subjects need to be categorized, no medication, oral agents and insulin, etc. At least, the subjects should be classified insulin use group or non-use group. And further analysis about correlation between Q score and patients' metabolic control status should be done in both groups

We agree with the reviewer. The medication history is given in the Methods section under the subheading “Patient data”. We have included a new table showing the patient characteristics (Table 1) and a new table showing the Q-Score and the Q-Score parameters (Table 3) for the following treatment groups: diet (no other medication), oral hypoglycaemic agents (OHA), OHA plus insulin and insulin alone. Figure 2 shows the relationships of the Q-Score with
diabetes therapy and the distribution according the treatment groups (diet, oral hypoglycaemic agents, oral hypoglycaemic agents plus insulin and insulin alone).

**Additional References**


