Title: Prediction models for short children born small for gestational age (SGA) covering the total growth phase. Analyses based on data from KIGS (Pfizer International Growth Database)

Authors:

Michael B Ranke (Michael.Ranke@med.uni-tuebingen.de)
Anders Lindberg (anders.lindberg@pfizer.com)

Version: 3 Date: 14 March 2011

Author's response to reviews: see over
Michael B. Ranke – Anders Lindberg

Dear Editor

BioMed Central

Thank you for giving us the opportunity to respond to the comments of the reviewers and to revise the article accordingly. We hope that the MS is now suitable for publication.

Sincerely yours
Reply to reviewers (point-to-point rebuttal)

We thank the reviewers for their constructive criticism. We have attempted to answer to the queries and to incorporate aspects whenever relevant. The reply to the points raised is set in italics following the original text of the query. Modifications in text of the MS were set in blue and underlined.

Reviewer 1: Maria de Ridder

- Major Compulsory Revisions

Point raised:
In an evaluation of prediction models in new, independent data, there should be given more information than only the mean of the residuals. It would be very interesting to see the performance (R2) of het model in the new group and whether the relation between predicted and observed values follows the line of identity. The latter can also be shown by checking whether the distribution of residuals is unrelated to the predictive values, as done in Figure 1e. It should be tested whether the decreasing pattern in this plot is significant.

Replic:
- The only new model in this paper is 3rd year prepubertal SGA.
- In table 2 there are mean and SD for the residuals (last row).
- There are no significant correlations in the residual plots 1a-1d.
- In figure 1e we've combined boys and girls (are actually based on two separate models). This is published in more detail elsewhere (ref 34). Combined there are a slightly negative correlation -0.27 (p=0.04), but since there are two different models the correlations are -0.10 (NS) and -0.14 (NS) for girls and boys respectively.
- Minor Essential Revisions

Abstract is lacking. – *This was a result of transferral. Is now added.*

The structure of Methods paragraph is not clear, partially due to formatting of the headings. - *The sub-headings were set in italics.*

Results:

All reported 'predicted height velocities' have the wrong unit. - *Yes, wrong units. Should be predicted height velocity in cm/y (not SDS as stated).*

Results of response first year:
It is not clear which test is reported in the last sentence. To compare predicted HV of the modelling and the evaluation group would be irrelevant. The subjects in the evaluation group must meet the criteria to apply the model, but mean values do not have to be comparable.

Why is noted p < 0.05 here, in contrast to in the Statistical paragraph: level p < 0.01? - *This is correct and leads to problems. We have decided to omit the last part of the sentence in the paragraph (... which is not significantly...) and refer correctly to Table 2.*

There are several inconsistencies between values in tables and text: - *This is correct, we have made the following changes:*

2nd prepub year:
- Change last sentence; SR of 0.04 (+/- 1.3) to *SR of 0.04 (+/- 1.2)*

3rd prepub year:
- Change in the middle; height of -3.7 (+/- 1.9) to ... (+/- 1.1)

4th prepub year:
- Change 22% boys to 62%
- Change in the middle; height of -4.2 (+/- 1.1) to ... (+/- 1.3)
In the text there is not any reference to the figures. – correct. The text included after the Results section is: The difference between observed and predicted height velocity during the first to fourth prepubertal year and of observed and predicted total growth during puberty expressed in terms Studentized Residuals are illustrated in Figures 1 a-e.

It is not clear which is the correct reference for the TPG-model: is it [5] or [34]? - 34 is the correct ref for the latest TPG models. As mentions above the TPG models are described in more detail in ref 34. Was changed in text.

- Discretionary Revisions

Compared to the descriptives given for the "first 4 year" groups, the information about the TPG-group is very limited.

It would be respectable if Table 3 is presented as an overview of the now available models, with references to the papers in which they are published before. One of the contributions added by the present paper could be to state the R2 and Error SD that are obtained now by using new data.

– Sorry, the authors are at variance with the reviewer.

Discussion:

How is the comparison between fourth year growth and spontaneous normal growth done?

– We presume that the reviewer wants to point out that there is little gain in height during the fourth year on GH in SGA. Even though this is frequently so we and other experts tend to believe that most of gain towards normalisation occurs during 4 prepubertal years as in other growth disorders.

Presenting all plots using the same scales for the axes is not desirable in this case, while the range of HV in succesive years is nog comparable.

- This choice was intended. Certainly it could be done otherwise.
Reviewer 2: Alan Rogol

Thank you for the comments. We have considered a minor suggestion.

Reviewer 3: Giorgio Bedogni

Major Compulsory Revisions

1. Please, explain better how missing data were handled. For instance, model 1 of table 3 was developed on 613 children and this is confirmed by the N = 613 in Table 1. However, potential predictors are not available in the same quantity for all children, e.g. n = 276 for height velocity. How did you handle the missing data? Did you consider ALL predictors (as apparently stated under Methods) or a subset of them? Did you do some data imputation? Please, make this clear for all the models. Also, consider reporting data ONLY for relevant subjects. –

No imputations were used. Observations with missing data was not used in any models.

Height velocity (pre GH therapy) N=276/N=183 was not used in 1st/2nd years prediction models. For clarity we therefore decided to delete the line from Figure 1.

2. Some "potential" predictors may be collinear. Did you check the presence of multicollinearity? How did you handle it if present?

- Yes, all standard co-linearity tests were performed when developing the prediction models. If present we selected the best combination of variables without any significant co-linearity (i.e. we could not include both height SDS and weight SDS).

3. All the models in table 3 assume a linear relationship between each predictor and the outcome. Did you check that all these relationships were linear? There is much to be gained in terms of accuracy if nonlinear relationships are modeled as such if present. Splines or fractional polynomials could be used to do that.

- Yes, we did check for non-linearity and transformed some of the variables (i.e. for IGHD models; max GH peak and GH dose). In the SGA models there was only linear relationships used. We have also in the
past modelling tested nonlinear modelling a lot without significantly improving the overall $R^2$ or reducing the predictive error level.

4. The number of subjects available for model 4 of Table 3 (4th year) appears low as compared to the number of potential predictors listed under Methods (see also comment #1).

The authors are not quite sure whether they understand the point raised. The numbers reported are the numbers which were available for the analysis. Certainly, there are reasons for the decrease in numbers: It is difficult possible to obtain complete longitudinal data on four years prepubertal growth since a) children - in particular girls - enter puberty, b) SGA is a rather "new" and not so frequent indication for GH (2002/2003) and c) many of the patients have not reached year 4 yet.

Minor Essential Revisions

1. Abstract: I suggest to replace "robust" with an other adjective, i.e. "accurate". You do not appear to have applied "robust" estimators in your analysis.

- Ok, change "robust" to "accurate".

2. I suppose that error SD in Table 3 is a root mean squared error of the estimate. Am I right?

- Yes, it is.

3. How was the "rank" in Table 3 calculated? Standardized regression coefficient? Other? Please, report the standardized regression coefficient instead if this is its ranking.

- We've used the "rank of importance" as an easier overview to compare what's most important predictors in the different models.
4. I suppose that the children who did not continue GH treatment over a given year did so for clinical reasons. I am right? Or someone exited the study for other reasons?

- The reasons why a patient did not continue treatment was decided by the investigator and is not always reported to KIGS (sometimes just lost to follow up).

5. Do you have any suggestion on how the accuracy of the models could be improved, e.g. by using other predictors? This may be useful for other researchers in the field.

- This is a principally interesting aspect, but beyond the scope of this article. It is discussed by the authors elsewhere (see refs)