Reviewer’s report

Title: Growth from birth to two years in a cohort of children diagnosed with CF by newborn screening

Version: 0 Date: 09 Apr 2019

Reviewer: Zhumin Zhang

Reviewer's report:

This study was designed to predict growth in the first two years of life using classification models. Identifying factors predicting growth and assessing their relative importance are of great to CF clinicians and researchers. But the manuscript has major issues related to outcome variables selected and relevant factors examined. Specific comments are included below for consideration.

1. The title "Growth from birth to two years in a cohort of children diagnosed with CF by newborn screening" is too broad, and did not reflect the main objectives of the study. In fact, growth patterns from birth to two years were not the main focus of the study.

2. Page 5 line 87 - "the remaining 129 children": the number seems not right (144 - 5 - 2 - 6 =131).

3. Page 5 line 95 - feeding type: please clarify that "breast" refers to exclusive breastfeeding.

4. Page 6 Line 115-116 & line 115-116: The four outcomes were the mean z-scores for height and weight in the first and second year of life. Please provide justifications for using the mean values of the entire year, instead of the actual measurements taken at age 1 and 2 years, respectively. As we all know, growth during infancy is dynamic and non-linear. For example, a healthy infant's birth weight doubles by age 6 months, and triples by age 1 year. The average values will not capture such dynamic changes. Also, please explain why growth indexes weight-for-length and BMI are not examined.

5. Page 6 line 104 - 105: Please explain why these cutoff points "-2, -1, and 0" were chosen. Moreover, not every child would "achieve" these values. What would be the time to achieve these cutoff values for these children?

6. Page 7 Line 125-130: Please report some of the characteristics of the study population, i.e., phenotype distribution (MI, PI, nonMI-PI), percent of infants born prematurely, percent of infants with low birth weight, genotype distribution, age at the first clinic visit. Also, the number/percentage of heterozygous for deltaF508 seems wrong (122/129).

7. Pancreatic status changes over time. Please specify when pancreatic status was assessed.

8. Page 7 Line 133-135: This statement is a bit misleading, as the test is not statistically significant. In fact, boys weigh more at birth than girls by ~100 gram in healthy, non-CF population (WHO growth reference).

9. Page 8 Line 136: Please specify exclusive breastfeeding if it is the case.

10. Appendix: it would be interesting to compare these two clusters with respect to distributions of
phenotype, genotype, gender, % of premature infants, % of infants with low birth weight, age at the first clinic visit, age at the CF diagnosis, etc.

11. Page 8 Line 155-157: Some of the variables change over time, i.e., fecal elastase, infant feeding, respiratory infections. Please specify when these variables were assessed, and explain how they were defined. The classification models could be biased without considering the changes of these variables during infancy.

12. Page 8 line 161: please define "poor infant growth".

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

Yes

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

No

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

No

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

Quality of written English
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published
Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

None

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal.