Reviewer’s report

**Title:** A new growth chart for preterm babies: Babson and Benda’s chart updated with recent data and a new format

**Version:** 3  **Date:** 7 October 2003

**Reviewer:** Marek Brabec

**Reviewer’s report:**

**General comments:**

The author selected an issue which is both practically important and difficult to deal with. The attempt to update older standards as well as to increase scope of the updated version is to be applauded. On the other hand, it is not only the compilation of the standard, but also some quality checking that makes the standard respectable and useful. Although the author notes right in the abstract that a validation “is needed to confirm that this growth chart is useful …”, no attempt to check the validity is attempted in this paper (apart from the comparison with the version being updated). This is the weakest point: without going through some checks, it is hard to decide whether the results are useful or not. The fact that the proposed standard was obtained via meta-analytic approach, which is known not to be problem-free in many cases (the substantial issue here is how representative the selected literature data are of the current preterm babies population), makes checking and substantial diagnostics of the proposed standards even more important. On the other hand, this is not to undermine the proposed standard per se. It might prove to be a very valuable one after going through the checks, but without them, it is hard to say, whether it is valuable or not.

**Compulsory revisions:**

1. **Page 2**
   “A validation study is needed to confirm that this growth chart is useful for monitoring …”
   This is very much true, without checking, it is hard to either criticize or praise the paper results. Even that thorough checking might take months to years, it would be nice to have results of at least some tests. Even small dataset(s) might help (either from one or a few facilities, and/or taken from another literature sources). One simple thing to look at would be the percentage of the babies below 10th percentiles and similar characteristics.
   Without at least some testing, it is hard to appreciate, whether the statements like “the larger sample sizes and more accurate gestational age assignments used here provide better confidence in the extremes.” (on page 2 and 13) are really true or whether they are products of wishful thinking.

2. **Page 4**
   “Three recentvery …”
   typo

3. **Page 5**
   “To develop the head circumference and length curves, the numerical data from … were averaged together using a weighted average based on total sample size.”
   This is certainly appropriate. But what about the measurement precision for different studies used? Have you checked, whether the precision was constant (using the same instruments and
procedures) or whether it changed from study to study? The paper does not say anything about is. In case that the precision would be different, it would be appropriate to weight studies inversely proportionally. Strictly speaking, a standard should relate to a particular measurement technique.

4. Page 6
“Since the differences in size between males and females … are not large, the data for both genders were averaged together …”
The fact that the inter-sexual differences are not large should be established more formally (statistical testing would be ideal), perhaps quoting some data, showing their numerical values, etc.

5. Page 6
“Smoothing of curves”
Although some reasons for smoothing are stated in this section, no details of how the smoothing was done are provided. In particular, information about how much smoothing was used (as measured e.g. by moving average window size, roughness penalty for splines, etc.) is missing. This is important, since without this info, it is not clear, how truthful is the constructed standard to the original data.

Page 7
“... adjustment was made for multiple comparisons.”
Nothing is said about how the adjustment was done. Many methods exist, that differ vastly in their conservativeness (ability to find real differences).

Page 9
“lengthand”
typo

Page 12
“realisticexpected”
typo

Discretionary revisions:
1. Page 2
“A literature search was conducted from 1980 to 2002 …”
Table 1
Page 2
“The new data produced curves that generally followed similar patterns as the old chart, with the difference …”
When creating the new standard and comparing it to the Babson’s old one, have you taken quite large time span of the studies used for the new standard construction into account? There might be a secular trend (as auxologists call it), that is the change in the growth tempo and/or attained size for the same age throughout calendar years. (Possibility of the secular trend is probably one of the reasons motivating the need for a new chart construction.) So that it is not completely correct to compare directly (without appropriate statistical adjustment) measurements taken at different calendar years.

Level of interest: An article whose findings are important to those with closely related research interests
Quality of written English: Acceptable

Declaration of competing interests:
None.