Reviewer's report

Title: Bioelectrical impedance analysis to estimate body composition in overweight and obese adolescents: comparison with dual-energy x-ray absorptiometry

Version: 4  Date: 15 August 2014

Reviewer: Lars Jødal

Reviewer's report:

The article "Bioelectrical impedance analysis to estimate body composition in overweight and obese adolescents: comparison with dual-energy x-ray absorptiometry" aims to test the ability of a given BIA device (Tanita BIA8) to estimate FFM, FM, and %BF in obese adolescents, using DXA as reference methods. As a secondary aim, the article also aims to assess %BF change after weight loss intervention.

Similar comparisons have been made before in other study populations, but the authors may well be the first to make such comparison in a population of overweight and obese adolescents.

As noted in the paper, a range of BIA devices exits. This reviewer has not specific knowledge of the Tanita BIA8 device, but using a multi-frequency, hand-to-foot, several (eight) electrodes BIA device seems like a good choice (rather than being limited to only one frequency or e.g. foot-to-foot). Especially since the authors compare results from multi-frequency data with results from using only one frequency.

The aims are clearly stated. Regarding the first aim, the authors fulfil their comparison, finding that all BIA methods tested have large variation, and that the value of multi-frequency BIA over single-frequency (50 kHz) BIA may be limited. The second aim (%BF change) is fulfilled, too, although given less attention.

As described in the numbered notes below, this referee has some questions for the methodology of the study, see especially notes 1 and 2. If these are resolved satisfactorily, the study is basically sound (with some questions are raised about more minor points).

The article needs strengthening on some points as described below, but is overall well-structured, and has a suitable headline. The literature seem well covered.

MAJOR COMPULSORY REVISIONS

1. During measurement of bioimpedance the subjects had arms "placed straight down by their sides" (line 153). Placing the arms in close contact with the side of the body would seem to involve a risk of current bypassing (part of) the arms,
following instead a more direct route to the rest of the body, especially if skin-to-skin contact was a possibility. In case this happens, measured resistance (and impedance) will be too low, which could lead to errors on the data level. It is therefore necessary that the authors address this issue.

2. In the discussion is noted that "DXA ... had been reported to overestimate adiposity by more than 20% in obese individuals.[22, 23]". This is a very important point, and credits go to the authors for bringing it up. But it seems to me that the implications are only partly discussed. Specifically, if DXA may overestimate obesity by 20%, then how can we know that the bias between DXA and BIA8 does not reflect BIA8 giving more correct values than DXA???. Referring to table 2, DXA gave FM on average 41.8 kg. Hypothetically, if this number is 20% too high, the true average FM would be 41.8/1.2 = 34.8 kg, in which case the Tatiana BIA8 would be the most correct (average) measure in the table. Please discuss further these implications for the study.

MINOR ESSENTIAL REVISIONS

3. The signs of differences are not consistent. For instance in the abstract, line 64-65: "BIA8 manufacturers equations ... overestimated FFM (4.3 kg[-13.9 to 5.3])". Here 4.3 kg is the overestimation (BIA8 - DXA), while the brackets are 95% CI for the difference with opposite sign (DXA - BIA8) - these signs are only clear when comparing with figure 1. Similar issues are found in the main text, e.g. lines 251-252, where the term "mean difference" is reported as a positive number for both overestimation and underestimation. (A discretionary suggestion on this issue is given in note 9 below.)

4. For several reasons, more details should be given about the clothing status of the investigated subjects.
   a) Metal in clothing may influence DXA measurements because metal is relatively impenetrable by the X-rays.
   b) Metal might also interfere with high-frequency currents in the BIA measurements.
   c) Clothing can be part of the answer to point 1 above (on bypassing currents).
   d) The precise weight of a subject will also depend on the amount of clothing worn during weighing.
   e) It is hard to find more information about the procedure used. Line 139 refers to "standard procedures previously described [7]", but in reference 7 this is only described by a very similar phrase "Weight, height and waist were measured using standard procedures previously described (24)".

5. In the derivation of equations, the participants are split in two groups for cross-validation (lines 198-202). As I understand the description, the cross-validation consists of comparing the predictive power of the equations derived from two groups (lines 214-216). However, given that the two groups
were random (and consistent) samples from the SAME initial population, it is no surprise that they have "effectively identical predictive power". Such splitting can, however, be used for estimating the uncertainties on the determined values, by comparing the results from the different sub-populations.

The authors should either omit the split in two groups, or use it only for calculations that can be expected to vary between the groups. (A reference for possible inspiration is given in note 13 below.)

6. In lines 291-292 the authors write that "There were no significant differences between the estimates of FFM from derived equations and DXA FFM for any of the models". This is not surprising: The model equations were BASED on the DXA data.

In general, if a model is based on a set of data, calculation of bias can only be used to evaluate the TYPE of model - a significant bias will signify that the model type cannot be applied to the data, even with the best-fit parameters. But bias from the original data cannot be used to evaluate the SPECIFIC model, since the model is a best-fit to the same data. The authors should therefore make clear that the bias in the models is not an independent measure - or omit reporting bias for the derived models.

7. Figure 2 shows compares DXA and BIA8 results for change in %BF over time. A "strong positive association" (line 305) is found in the plot. As I read the manuscript, the comments to this finding is given in lines 339-341, but different wordings are used in line 305 and lines 339-341, which obscures the connection. Please strengthen the connection between this finding and the discussion of the finding, e.g. by use of more similar wording.

8. The text on the y-axis (vertical axis) in Figures 2a and 2b does not make sense. Please correct.

DISCRETIONARY REVISIONS

9. Suggestion regarding note 3: This reviewer personally finds it most intuitive to describe a difference as (tested value - reference value). With this convention, overestimation will have a positive sign, while underestimation will have a negative sign. Lines 65-66 in the abstract could then be expressed along these lines:

"... overestimated FFM (mean difference 4.3 kg [95% CI, -5.3 to 13.9 kg]) and underestimated %BF (-5.0% [-15.1 to 5.0]) compared to DXA."

This formulation has consistent signs in all numbers (+ means too high, - means too low, compared to reference), and also makes clear that the brackets denote 95% CI.

10. If possible, it would be nice have an indication of how the BIA8 calculates FFM, FM and %BF. Or are these calculations completely "black-box"?

11. In lines 223-226, a test for outliers at the 0.05 level is made. Two points are classified as outliers and investigated. Since no problems are identified with the
points, they are retained. The authors may wish to include a sentence along this line: "With 66 subjects in the study, a test working at the 0.05 level will be expected to find in the order 0.05 x 66 ~ 2 'outliers' in a sound dataset" (as a further indication that it is a sound decision to retain the two 'outliers' in the dataset).

12. It is suggested to number the equations (1), (2), etc.

13. An example of cross-validation by sub-sampling can be found in Andersen et al, Am J Kidney Dis. 59(1):50-57.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests.