Reviewer's report

Title: Comparison of air displacement plethysmography to hydrostatic weighing for estimating total body density in children

Version: 1 Date: 9 May 2005

Reviewer: Randy Clark

Reviewer's report:

General

-------------------------------------------------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

The study examined the accuracy, precision and systematic bias of body density and percent body fat by comparing air displacement plethysmography (ADP) to hydrostatic weighing (HW) in 40 male and 26 female children.

Accuracy, precision, and bias of ADP were examined for the group and by gender with HW (with simultaneous measurement of residual lung volume by using the closed-circuit oxygen dilution technique) serving as the criterion method. Lohmans age and gender equations were use to convert body density to % fat. Two critical values in both procedures, thoracic gas volume (TGV) and residual lung volume (RV) were measured and not predicted. However, the RV value was acquired during submersion. This likely made an already difficult procedure even more complex (and perhaps uncomfortable) for the subjects. Maturation was not assessed. The subjects were fit and lean children recruited from a youth soccer league.

Evaluating new methodology such as ADP is particularly warranted in youth. An alarming 31% of children aged 6 to 19 years old are found to be at risk for overweight while 16% were classified as overweight. This is the first time in history that today's children will have a shorter life span than their parents. Over 10 million children are considered inactive and overweight. This increase the risk for type II diabetes, high cholesterol, high blood pressure, heart disease, cancer and psychological problems at a young age. From a medical standpoint this has been described by World Health Organization as one of the most important public health threats because of the significant long-term health consequences. Therefore, the evaluation of ADP is particularly important in this age group and the study is warranted. However, youth pose a special challenge to any research team because measuring body composition in children is particularly challenging and difficult.

The authors evaluate accuracy, precision and systematic bias using HW as their criterion method as follows:

Accuracy

The authors conclude that ADP is accurate when the regression between body density and percent fat by HW and ADP did not have a slope significantly different from one and an intercept significantly different from zero. This tested their hypothesis that the regression of body density and percent fat by HW and body density and percent fat by ADP did not significantly deviate from the line of identity.
Precision
The authors offer R^2 and SEE values as their evaluation of precision.

Bias
The authors use a Bland Altman analysis to evaluate bias between ADP and HW. The authors suggest that no bias exists between the techniques if there is a nonsignificant P value from the regression line that fits the residual plot.

Their methodology and analysis follow standard procedures. The manuscript is well written and flows well. The analysis and figures are consistent with previous work accepted in the body composition literature and closely mimic previous work by Fields and colleagues. The writing style paraphrases and, at times, duplicates word-for-word the previous work by Fields and Goran (Fields, D.A. and M.I. Goran, Body composition techniques and the four-compartment model in children. J Appl Physiol, 2000. 89(2): p. 613-20). However, the criterion in the previous study was the more complete four compartment model. The limitations of the methodology and statistical analysis are appropriately acknowledged.

Interestingly, the authors conclude a significant gender difference was detected when assessing total body density and % fat in males, however no difference was observed in the females. They describe the results from this study as provocative. This requires clarification.

The following concerns require attention or clarification:

Simultaneous measurement of RV. Relatively untested in this population and a difficult procedure for any subject, let alone children. Requires justification and clarification.

The authors suggest The TGV is analogous to the residual volume in HW. These are very different values, measured in different ways. This is not accurate and requires correction. Clarify the difference and how these values are used in calculation of BD by each method.

Most concerning is that the statistical analysis of the data do not support the gender differences concluded by the authors. For example, the authors correctly acknowledge that a power analysis for this data set was performed and it was found after factoring in those dropped from the study, there remained only a power of 0.429. Therefore, we did not have enough power to detect the potential difference between the methods in females. Yet having made this statement the study’s main finding is a gender difference between the methods. There is not sufficient power in females to make this claim.

The coefficient of variance (CV) for repeated measures over two days in a subset of the children in this study for ADP and HW was 3.1% and 7.1% respectively. This is high when comparing the CV for ADP and HW in other laboratories. In fact the authors go on to say that more typical CVs in their laboratory for adults are 1% and 1.5% respectively. The authors speculate one possible reason for the differences found in CV for adults and children could have been due to poor subject compliance resulting in invalid measurements. However, speculation that the inability of the children to comply with the requirements to complete a valid HW test may be at the root of the differences found in this study is a very curious statement indeed. Particularly when the authors judge the accuracy, precision and bias of the ADP technique against their HW criterion value.

Why are these results provocative?

The group (Y = 0.835x + 0.171, R^2=0.84, SEE=0.007 g/cm3) and male regressions (Y = 0.837x + 0.174, R^2=0.90, SEE=0.006 g/cm3) of ADP total body density compared to HW total body density significantly deviated from the line of identity.
Precision?
The regression between % fat by HW and % fat by ADP significantly deviated from the line of identity for the entire group ($Y = 0.84x + 3.81, \text{R}^2=0.83, \text{SEE}=3.35 \% \text{fat}$) and in the males ($Y = 0.84x + 3.25, \text{R}^2=0.90, \text{SEE}=3.00 \% \text{fat}$). However, in females the regression did not significantly deviated from the line of identity ($Y = 0.81x + 5.17, \text{R}^2=0.56, \text{SEE}=3.80 \% \text{fat}$). Please discuss the meaning and importance of a larger SEE even though did not deviate from line of identity. The authors base precision on R2 and SEE, however this is not discussed. Requires elaboration and clarification. How is R2 used as a measure of precision?

A portion of the error can be contributed to the criterion. Particularly based on author comments above and simultaneous measurement of RV. Elaborate on the contribution of error in the HW criterion measure.

The authors correctly make an attempt to eliminate 4 of the common problems in ADP standardization:

- Measured vs. predicted TGV
- Strict clothing protocol
- Child specific equation when converting body density to % fat,
- ADP preceded HW

Please discuss the effects of body temperature and body surface area in these children when using ADP in this study.

Clearly a 4C model would be preferred in this sample, such as previous work by Fields and Goran compared DXA, HW and ADP to the four compartment model in children aged 9-14 years old. Is this a subset of previous 4 compartment data or new work?

The four-compartment model is now considered the research standard criterion method to assess body composition because the individual constituents of the fat free mass are measured. However, due to expensive equipment and testing procedures it is not commonly used by most laboratories. This does not rule out HW as a criterion and a comparison between HW and ADP is valuable. However the research team needs to be confident in the HW measures. This does not seem to be the case in this study.

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: Yes

Declaration of competing interests:

I declare that I have no competing interests