**Team C**

**General Comments**

1. The authors explained thoroughly the quadratic models fitted (for Aim 1a) in the Results section. However, part of that explanation seems to belong to the Methods.
2. Part of the Discussion could be better fit in the Results section (e.g. all the results on gender).
3. The authors use the word “diagnostic” in two different contexts: when the validity of model assumptions and when referring to which criteria to use to pick the best models.
4. It is stated in the Discussion that external validation was not able to be carried out. Why is this? Will there be an external validation carried out and discussed in the final report?
5. Using *Aim 1a*, for example, as a section heading is not informative. Add a few words to remind the reader about aim of *Aim 1a*.

**Specific Comments**

**Abstract**

1. [Objective] Use complete sentences.
2. [Methods] What best model selection did the authors used? Do the authors refer to the AIC specified in the results section?
3. [Methods] “Forward and backward stepwise selection” is not used to “predict gender”, but rather to build a model relative to a procedure, such as logistic regression.
4. [Results] What does an AIC value of -1482.04 reveal? Note that AIC is a relative measure used to compare different models.
5. [Results] What are the results of the internal validations?

**Background**

1. [First paragraph] The layman will not know what “constant measures” means. What is constant about these measures? What is nonconstant about the other measures?
2. [First paragraph] What “conflicting results” were presented? Do you mean for the last sentence to answer this?
3. [Second paragraph] The statement on BMI is not quite correct. BMI accounts for both fat mass and muscle mass (as both those measures affect weight). The problem with BMI is the inability to distinguish fat mass from lean and muscle mass.
4. [Third paragraph, last sentence] Are these relationships not well known, or is it known that measures other than those on the pelvis are not as discriminating?

**Methods**

1. [First paragraph] This is picky, but a dataset contains measures, not participants.
2. [Aim 1a] It is unclear when the authors replaced skeletal ankle, knee and wrist with the corresponding girths measures. Were the 3 skeletal variables replaced in the model selected by the AIC?
3. [Aim 1b: first sentence] “We used the predicted weight from our model…” which model? The linear or the quadratic model? Although the authors made it clear in the results section which model they are using, there should be a reference on which model was used in the Methods.
4. [Aim 1b] How did the authors checked “the utility of BMI among subjects with high/low fat mass or muscle mass”?
5. [Aim 2] There is no need to use both AIC and BIC. The authors should pick one. Note that only AIC was mentioned in the Abstract.

**Results**

1. [First paragraph] Age and BMI units of measure are missing.
2. [Aim 1a: second paragraph] The first quadratic model is the one used by Benkhe et al. The authors should add a reference when explaining the model (reference #12).
3. [Aim 1b]: BMI unit of measure is missing.
4. [Aim 1b: last sentence] “…than males with low muscle mass.” There is no need to repeat the sentence.
5. [Aim 2] “…stepwise regression had best log-likelihood, AIC, and BIC”. What do the authors mean by *best*?

**Discussion**

1. [Second paragraph] How do the authors defined a “reasonable amount” of influence and leverage points.
2. [Third paragraph] “…we failed to reject the null” which null do the authors refer to? What difference was observed? The large p-value suggests that the observed difference might be due to chance and not a true difference in the population.

**Tables and Graphs**

1. [Table 3]: BMI unit of measure is missing.