Supplemental Digital Content 1

Exploration of scaling functions

Introduction

A reviewer suggested that we rigorously evaluate a number of suggested size scaling functions. The goal being to determine whether the allometric scaling based on predicted fat-free-mass (FFM) by the Al-sallami equation[1] is optimal for the PK data. Their suggested approach was to evaluate the first few steps in model development and identify the best performing model as the most promising direction for further model development. Unfortunately, our experience in model development is that the first few steps in model development are not always representative of the achievable performance for the final model. Scaling functions which perform best for the first few steps of model development often "stall" and other scaling functions can "overtake" them and achieve a better performing final model. Since we are interested in obtaining a good performing final model, and not necessarily a good performing initial model, to really rigorously evaluate size scaling functions we must perform a full model development for each size scaling function under consideration. We performed this for 6 size scaling functions suggested by the reviewer. These were:

1. Linear weight size scaling
2. Al-sallami FFM equation[1] size scaling
3. Janmahasatian FFM equation[2] size scaling
4. Dubois equation[3] surface area scaling
5. Weight power exponent scaling
6. Weight with BMI correction scaling

We display the hierarchical model development processes in figures along with a short summary. In the figures, models shaded blue are "final" models for each branch of model development. For these models the correlations between age, weight, height and BMI and posthoc eta for all parameters were $p>0.01$. This condition suggests that further model

development is not likely to be beneficial. Evaluating model modifications for inclusion was performed using the same criteria as in the manuscript. We required a decrease in AIC of at least 10 for adding parameters to the model and allowed an increase in AIC of up to 4 when removing parameters. We also considered predictive performance across 5 subgroups: young children (age < 3 yr), children (3 ≤ age <18 yr), young-adults (18 ≤ age < 40 yr), middle-aged (40 ≤ age < 65 yr) and elderly (age ≥ 65 yr). The overall measure of predictive performance was the average $MdAPE_{PK}$ and across these subgroups using all predictions from 10 repeats of randomly partitioned two-fold cross-validation. We did not consider models that show clear degradation of the ability of models to predict the observations, even if such models showed improved log-likelihood of fit to the data.

For all model development figures: AIC is the Akaike Information criteria, k is the number of model parameters, MdAPE is the median absolute prediction error for all observations and Avg MdAPE is the median (cross-validation) absolute prediction error averaged over the subgroups young-children, children, young-adults, middle-aged and elderly.

Caveat

We performed this exercise of comparing model development using several different scaling functions at the request of the reviewer. However, we would not suggest this practice in general for other investigations. We could think of hundreds of different potential scaling functions which have no biological basis and test them all. However this would increase the risk of false-positive results due to multiple-testing. We believe it is beneficial to require at least some biological justification to test a proposed size scaling function. At the very least there should be a simple explanation rooted in biology why the scaling function might be relevant. As an example, we would not have tested the Dubois surface area function as a size scaling function for volumes. For a pharmacological volume to scale with surface area requires that the pharmacologically active volume be a thin shell under the surface. We don't think this is realistic, so we would not have tested that size scaling function outside of the reviewers request to do so.

Linear weight size scaling

Here the size function scales linearly with weight. The size scaling equation is:

$$SIZE = \frac{WGT}{70.}$$

where *WGT* is weight in kg.

Hierarchical model development is shown in Figure 1. Here we briefly summarize the steps in model development.

- For the initial model (Model 1) all parameters scale linearly with SIZE.
- Adding aging to this model (Model 2) improved the model.
- The addition of maturation functions based in age (Model 3) or weight (Model 4) did not improve the model.
- Increasing V2, CL and Q2 for females aged 12-45 years (Model 5) did not improve the model, however simply increasing CL for females (Model 6) did improve the model.
- Model 6 shows a clear deviation in posthoc eta for CL and adding an exponential weight correction function to CL (Model 7) resulted in an improved model.

- Model 7 shows a clear deviation in posthoc eta for V2 and adding an exponential weight correction function to V2 (Model 8) resulted in an improved model.
- Model 8 is the final model for this branch in model development showing models all correlations between age, weight, height and BMI and posthoc eta for all parameters were $p>0.01$. The AIC was -6344.22 with an average MdAPE from cross-validation of 24.0%.
- Adding allometric scaling to the initial model (Model 9) resulted in an improved model.
- Adding aging to this model (Model 10) resulted in an improved model.
- Adding compartmental allometry (Model 11) to the allometric scaled model improved AIC by >100 but predictive performance remained the same.
- Adding aging to this model (Model 12) resulted in an improved model.
- The addition of maturation functions based in age (Model 13) improved the model but basing maturation function on weight (Model 14) improved the model to a greater degree.
- Increasing CL for females (Model 15) improved the model but increasing V2, CL and Q2 for females aged 12-45 years (Model 17) improve the model to a greater degree.
- Both Model 15 and Model 17 show correlation between weight and in posthoc eta for V3. Adding an exponential weight correction function for V3 improved the model in both cases. (Model 16 and Model 18).
- Model 16 and Model 18 are the final models for this branch in model development showing models all correlations between age, weight, height and BMI and posthoc eta for all parameters were $p>0.01$. The AIC was -6476.50 and -6501.85 with an average MdAPE from cross-validation of 22.9 and 21.3%.

The summarized findings of model development for linear weight scaling are:

1) Including allometric scaling and compartmental allometry resulted in the best performing final model, both in terms of AIC and lowest MdAPE for cross-validation.
2) Maturation functions are only beneficial when allometric scaling is used.

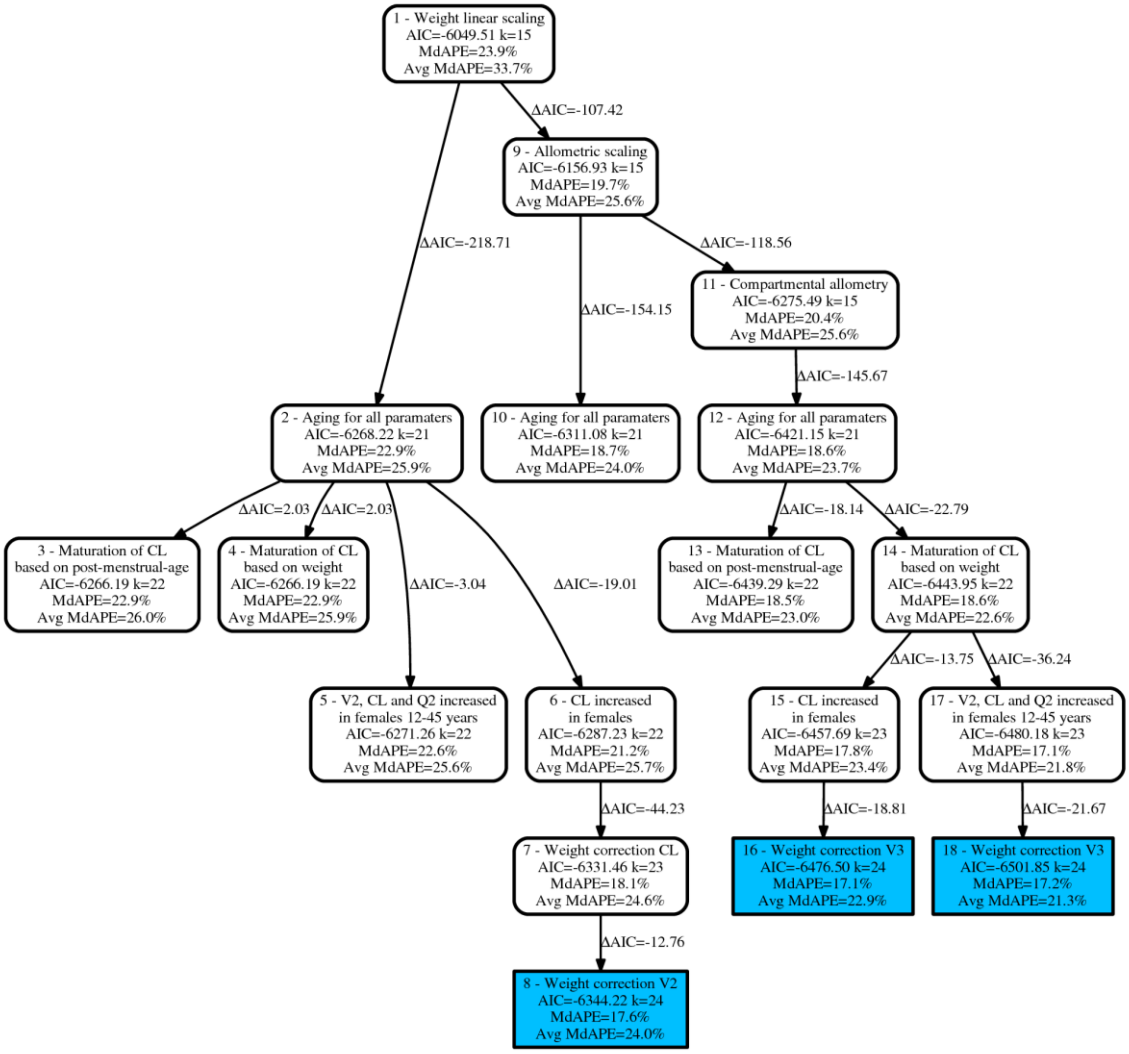

Figure 1. Hierarchical PK model building for size scaling linear with weight.

Al-sallami FFM size scaling

Here the size function scales FFM as calculated by the Al-sallami equation. The size scaling equation is:

$$SIZE = \frac{FFM}{FFM_{ref}}$$

Where FFM_{ref} is the FFM calculated for the reference individual, a 70-kg, 170-cm, male.

Hierarchical model development is shown in Figure 2. Here we briefly summarize the steps in model development.

- For the initial model (Model 1) all parameters scale linearly with FFM as defined by the Al-sallami equation.
- Adding aging to this model (Model 2) improved the model.
- The addition of maturation functions based in age (Model 3) or weight (Model 4) did not improve the model.
- Increasing CL for females (Model 5) improved the model and increasing V2, CL and

Q2 for females aged 12-45 years (Model 8) also improved the model.

- Model 5 showed differences in posthoc eta for V1, V2 and Q2 for females compared to males. Increasing these parameters for females (Model 6) improved the model.
- Model 6 show correlation between weight and in posthoc eta for CL. Adding an exponential weight correction function to CL (Model 7) improved the model.
- Model 6 show correlation between weight and in posthoc eta for CL. Adding an exponential weight correction function to CL (Model 7) improved the model.
- Model 7 is the final model for this branch in model development showing models all correlations between age, weight, height and BMI and posthoc eta for all parameters were $p>0.01$. The AIC was -6352.03 with an average MdAPE from cross-validation of 24.1%.
- Model 8 show correlation between weight and in posthoc eta for CL. Adding an exponential weight correction function to CL (Model 9) improved the model.
- Model 9 show correlation between weight and in posthoc eta for V2. Adding an exponential weight correction function to V2 (Model 10) improved the model.
- Model 10 is the final model for this branch in model development showing models all correlations between age, weight, height and BMI and posthoc eta for all parameters were $p>0.01$. The AIC was -6355.64 with an average MdAPE from cross-validation of 22.8%.
- Adding allometric scaling to the initial model (Model 11) resulted in an improved model.
- Adding aging to this model (Model 12) resulted in an improved model.
- Adding compartmental allometry (Model 13) to the allometric scaled model improved the model.
- Adding aging to this model (Model 14) resulted in a decrease in AIC of nearly 100 and predictive performance was slightly degraded (0.1%). We provisionally accept this step into the model and revisit simplifications to the aging model later in model development.
- The addition of maturation functions based in age (Model 15) improved the model but basing maturation function on weight (Model 16) improved the model to a greater degree.
- Increasing CL for females (Model 17) improved the model but increasing V2, CL and Q2 for females aged 12-45 years (Model 18) improve the model to a greater degree.
- Estimating the 12 and 45 year age boundaries (Model 19) did not result in an improved model.
- Model 18 show correlation between weight and in posthoc eta for V3. Adding an exponential weight correction function for V3 (Model 20) improved the model.
- Reexamination of the age correction functions showed the age correction function parameters for V1, Q2 and Q3 were very similar as were those from V2 and CL. These could be combined (Model 21) to reduce the number of parameters in the model. AIC decreased and predictive performance was unchanged.

- Model 21 is the final model for this branch in model development showing models all correlations between age, weight, height and BMI and posthoc eta for all parameters were p>0.01. The AIC was -6519.65 with an average MdAPE from cross-validation of 20.8%.

The summarized findings of model development for weight scaling using the Al-sallami FFM equation are:

1) Including allometric scaling and compartmental allometry resulted in the best performing final model, both in terms of AIC and lowest MdAPE for cross-validation.
2) Maturation functions are only beneficial when allometric scaling is used.

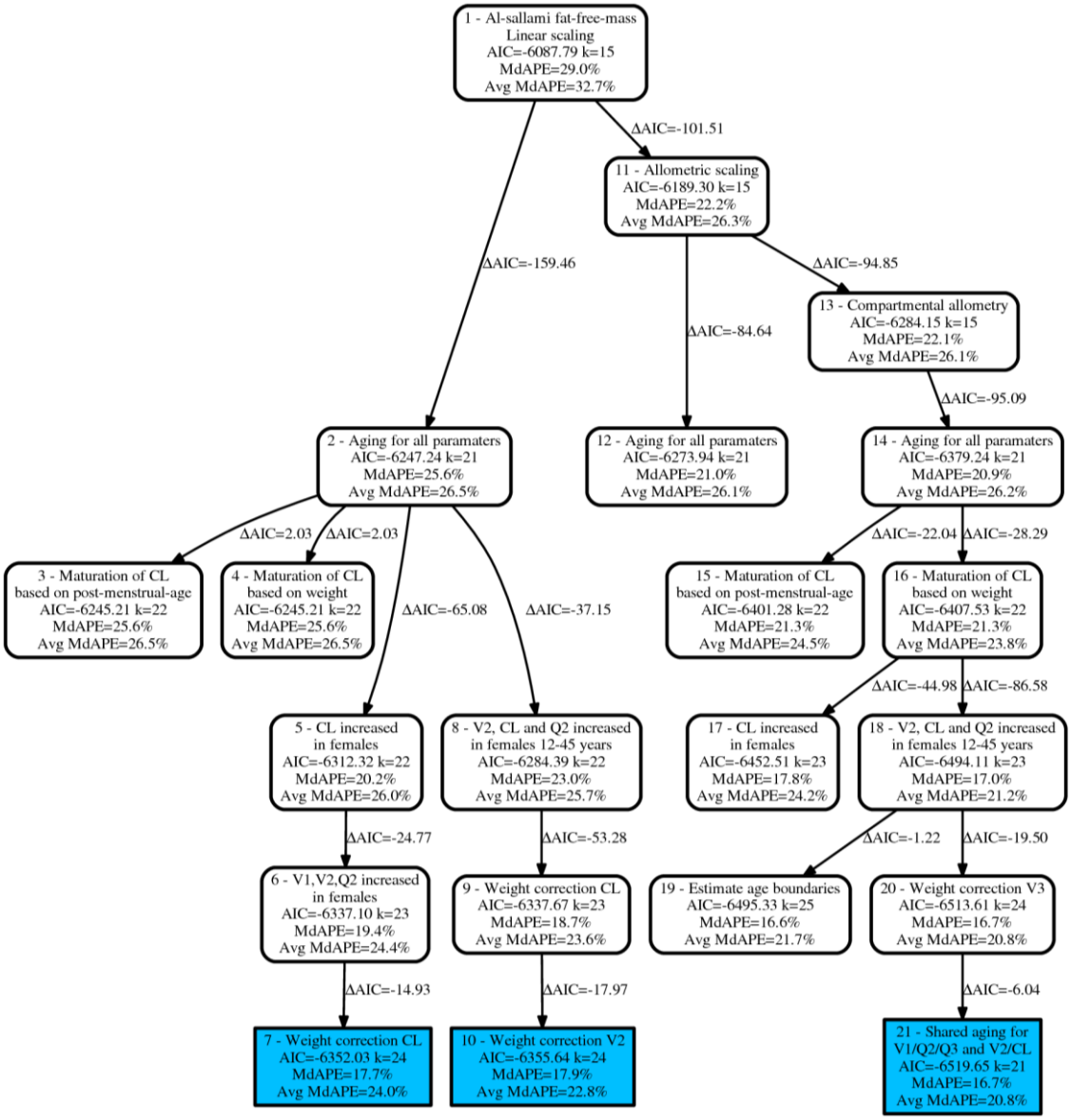

Figure 2. Hierarchical PK model building for size scaling using the Al-sallami FFM equation.

Janmahasatian FFM size scaling

Here the size function scales FFM as calculated by the Janmahasatian equation. The size scaling equation is:

$$SIZE = \frac{FFM}{FFM_{ref}}$$

where FFM_{ref} is the FFM calculated for the reference individual, a 70-kg, 170-cm, male.

It is difficult to justify use of the Janmahasatian FFM function while the Al-sallami FFM function is available. The Janmahasatian FFM function is biased for small children and the Al-sallami FFM equation corrects this bias. As expected the Janmahasatian FFM function performed very similarly to scaling with the Al-sallami FFM equation.

Hierarchical model development is shown in Figure 3. Here we briefly summarize the steps in model development.

- For the initial model (Model 1) all parameters scale linearly with FFM as defined by the Janmahasatian equation.
- Adding aging to this model (Model 2) improved the model.
- The addition of maturation functions based in age (Model 3) or weight (Model 4) did not improve the model.
- Increasing CL for females (Model 5) improved the model and increasing V2, CL and Q2 for females aged 12-45 years (Model 7) also improved the model.
- Model 5 showed differences in posthoc eta for V1, V2 and Q2 for females compared to males. Increasing these parameters for females (Model 6) improved the model.
- Model 6 is the final model for this branch in model development showing models all correlations between age, weight, height and BMI and posthoc eta for all parameters were $p>0.01$. The AIC was -6344.80 with an average MdAPE from cross-validation of 23.9%.
- Model 7 show correlation between weight and in posthoc eta for CL. Adding an exponential weight correction function to CL (Model 8) improved the model.
- Model 8 is the final model for this branch in model development showing models all correlations between age, weight, height and BMI and posthoc eta for all parameters were $p>0.01$. The AIC was -6327.99 with an average MdAPE from cross-validation of 23.5%.
- Adding allometric scaling to the initial model (Model 9) resulted in an improved model.
- Adding aging to this model (Model 10) resulted in an improved model.
- Adding compartmental allometry (Model 11) to the allometric scaled model improved the model.
- Adding aging to this model (Model 12) improved the model.
- The addition of maturation functions based in age (Model 13) improved the model but basing maturation function on weight (Model 14) improved the model to a greater degree.
- Increasing CL for females (Model 15) improved the model but increasing V2, CL and Q2 for females aged 12-45 years (Model 17) improve the model to a greater degree.
- Model 15 showed correlation between weight and in posthoc eta for V3. Adding an

exponential weight correction function for V3 (Model 16) improved the model.

- Model 16 is the final model for this branch in model development showing models all correlations between age, weight, height and BMI and posthoc eta for all parameters were p>0.01. The AIC was -6464.01 with an average MdAPE from cross-validation of 24.0%.
- Model 17 showed correlation between weight and in posthoc eta for V3. Adding an exponential weight correction function for V3 (Model 18) improved the model.
- Model 18 is the final model for this branch in model development showing models all correlations between age, weight, height and BMI and posthoc eta for all parameters were p>0.01. The AIC was -6507.35 with an average MdAPE from cross-validation of 21.0%.

The summarized findings of model development for weight scaling using the Janmahasatian FFM equation are:

1) Including allometric scaling and compartmental allometry resulted in the best performing final model, both in terms of AIC and lowest MdAPE for cross-validation.
2) Maturation functions are only beneficial when allometric scaling is used.
3) Model development is very similar to when the Al-sallami FFM equation is used for size scaling.

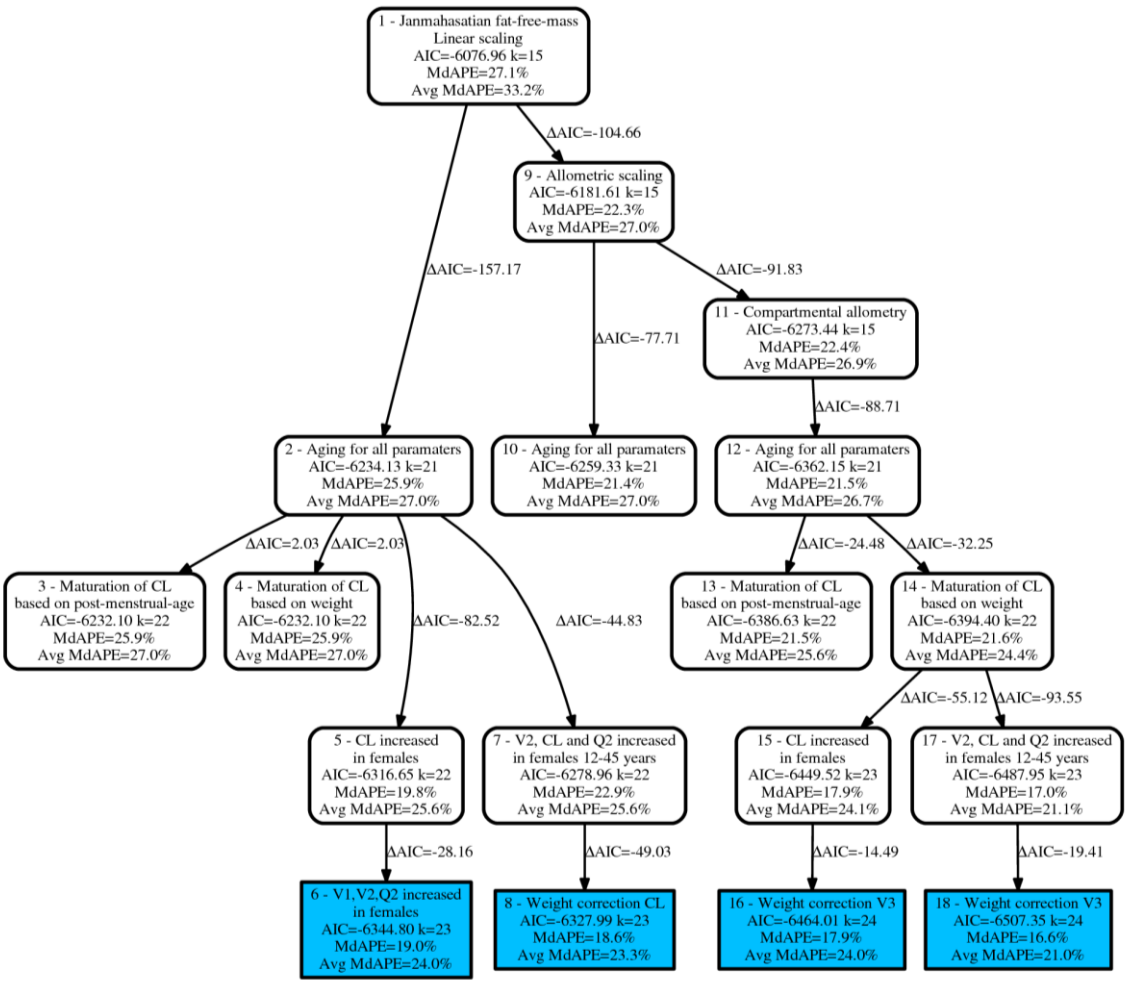

Figure 3. Hierarchical PK model building for size scaling using the Janmahasatian FFM equation.

Dubois equation surface area scaling

Here the size function scales as surface are calculated by the Dubois body surface area (BSA) function. The size scaling equation is:

$$Surface = 0.20247 \cdot HGT^{0.725} \cdot WGT^{0.425}$$

$$SIZE = \frac{Surface}{Surface_{ref}}$$

Where *HGT* is height in meters and *WGT* is weight in kg, and $Surface_{ref}$ is calculated for the reference individual, a 70-kg, 170-cm, male.

It is difficult justify the use of this model because the mechanistic interpretation for volumes is unclear. Volumes can scale with surface area if the pharmacologically active volume forms a thin shell under the surface and this seems unrealistic.

Hierarchical model development is shown in Figure 4. Here we briefly summarize the steps in model development.

- For the initial model (Model 1) all parameters scale linearly with surface area predicted by the Dubois equation.
- Adding aging to this model (Model 2) improved the model.
- The addition of maturation functions based in age (Model 3) improved the model but basing maturation function on weight (Model 4) improved the model to a greater degree.
- Increasing CL for females (Model 5) improved the model, however increasing V2, CL and Q2 for females aged 12-45 years (Model 6) improved the model to a greater degree.
- Model 5 is the final model for this branch in model development showing models all correlations between age, weight, height and BMI and posthoc eta for all parameters were p>0.01. The AIC was -6364.69 with an average MdAPE from cross-validation of 23.2%.
- Model 6 is the final model for this branch in model development showing models all correlations between age, weight, height and BMI and posthoc eta for all parameters were p>0.01. The AIC was -6377.10 with an average MdAPE from cross-validation of 21.5%.
- Adding allometric scaling to the initial model (Model 7) resulted did not result in an improved model.
- Adding aging to this model (Model 8) improved AIC but degraded predictive performance.
- Adding compartmental allometry (Model 9) to the allometric scaled model improved the model.
- Adding aging to this model (Model 10) resulted in an improved in AIC however predictive performance was degraded.
- The addition of maturation functions based in age (Model 11) improved the model but basing maturation function on weight (Model 12) improved the model to a greater degree.

- Increasing CL for females (Model 13) did not improve the model sufficiently to justify the additional model parameter. Increasing V2, CL and Q2 for females aged 12-45 years (Model 14) did improve the model.
- Model 14 is the final model for this branch in model development showing models all correlations between age, weight, height and BMI and posthoc eta for all parameters were p>0.01. The AIC was -6473.91 with an average MdAPE from cross-validation of 21.8%.

The summarized findings of model development for size scaling using the Dubois body surface area equation are:

1) With allometric scaling and compartmental allometry the final model had the best AIC (-6473.91 vs. -6377.10) however predictive performance in cross-validation was only slightly poorer (21.8% vs. 21.5%)
2) Maturation functions were beneficial both with and without allometric scaling.

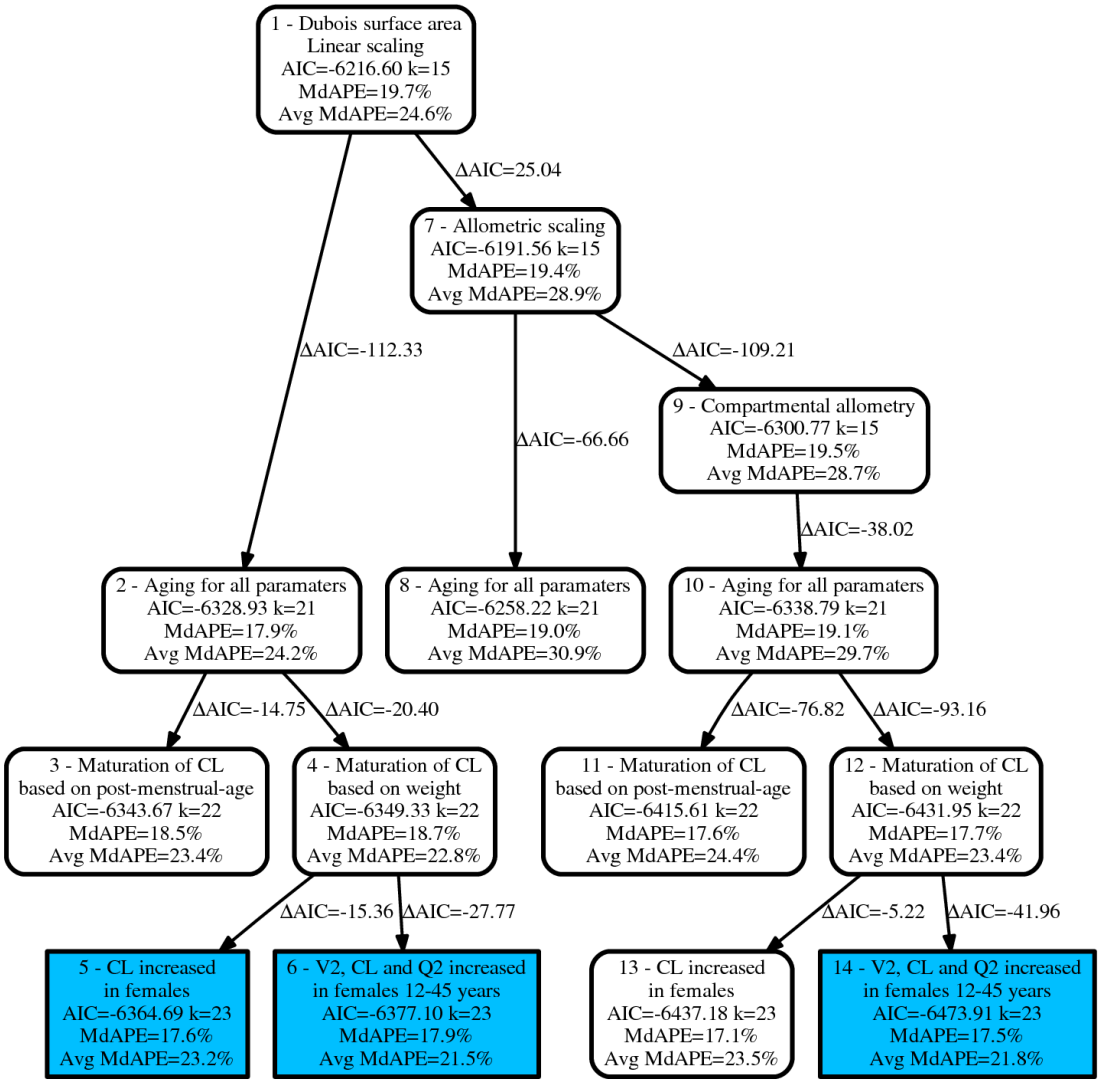

Figure 4. Hierarchical PK model building for size scaling using the Dubois formula for surface area.

Weight power exponent scaling

Here the size function scales with weight to a power which is estimated from the data. The

size scaling equation is:

$$SIZE = \frac{WGT^x}{70^x}$$

Where *WGT* is weight in kg and *x* is estimated from the data during model fitting.

Hierarchical model development is shown in Figure 5. Here we briefly summarize the steps in model development.

- For the initial model (Model 1) all parameters scale linearly with an estimated weight power exponent. The estimated exponent is about 0.67.
- Adding aging to this model (Model 2) improved the model.
- The addition of maturation functions based in age (Model 3) did not improve the model sufficiently to justify the additional parameter. However, basing maturation on weight (Model 4) did improve the model.
- Increasing CL for females (Model 5) improved AIC but degraded predictive performance in cross-validation. Increasing V2, CL and Q2 for females aged 12-45 years (Model 6) did improve the model.
- Model 6 is the final model for this branch in model development showing models all correlations between age, weight, height and BMI and posthoc eta for all parameters were $p>0.01$. The AIC was -6362.84 with an average MdAPE from cross-validation of 22.3%.
- Adding allometric scaling to the initial model (Model 7) did not improve the model.
- Adding aging to this model (Model 8) resulted in an improved model.
- Adding compartmental allometry (Model 9) to the allometric scaled model improved AIC but predictive performance in cross-validation was decreased.
- Adding aging to this model (Model 10) resulted in an improved model.
- The addition of maturation functions based in age (Model 11) improved the model but basing maturation function on weight (Model 12) improved the model to a greater degree.
- Increasing CL for females (Model 13) improved AIC but degraded predictive performance in cross-validation. However, increasing V2, CL and Q2 for females aged 12-45 years (Model 14) improved the model.
- Model 14 show correlation between weight and posthoc eta for V3. Adding an exponential weight correction function for V3 (Model 15) improved the model.
- Model 15 is the final model for this branch in model development showing models all correlations between age, weight, height and BMI and posthoc eta for all parameters were $p>0.01$. The AIC was -6504.04 with an average MdAPE from cross-validation of 21.2%.
- The estimated exponent for the best model (Model 15) is about 0.91.

The summarized findings of model development for weight scaling using weight power (estimated) exponent are:

1) Including allometric scaling and compartmental allometry resulted in the best performing final model, both in terms of AIC and lowest MdAPE for cross-validation.
2) Maturation functions based on weight were beneficial both with and without allometric scaling and compartmental allometry.
3) The estimated exponent for the initial model was about 0.67, closely approximating the 2/3 allometric exponent for surface area, however the estimated exponent for the best final model was about 0.91, which is much closer to linear scaling.

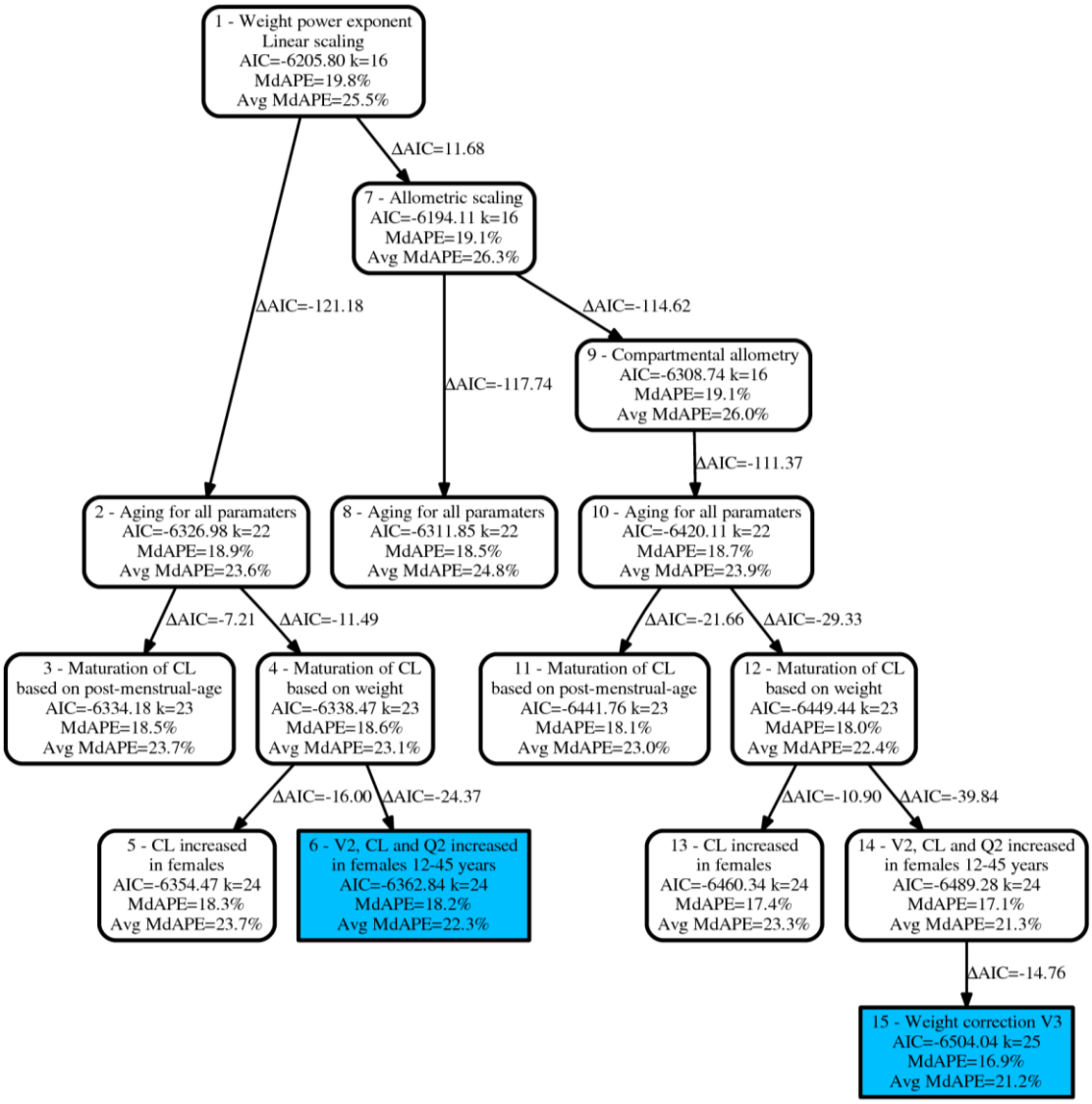

Figure 5. Hierarchical PK model building for size scaling using weight to an estimated power exponent.

Weight with BMI correction scaling

Here the size function scales size with weight corrected by BMI. The size scaling equation is:

$$SIZE = WGT \cdot \exp(x \cdot (BMI - BMI_{ref}))$$

Where WGT is weight in kg and x is estimated from the data during model fitting. BMI_{ref} is the reference BMI for a 70-kg, 170-cm individual.

Hierarchical model development is shown in Figure 6. Here we briefly summarize the steps in

model development.

- For the initial model (Model 1) all parameters scale linearly with an (estimated) BMI corrected weight.
- Adding aging to this model (Model 2) improved the model.
- The addition of maturation functions based in age (Model 3) or weight (Model 4) did not improve the model.
- Increasing CL for females (Model 5) improved AIC but degraded predictive performance in cross-validation. Increasing V2, CL and Q2 for females aged 12-45 years (Model 6) did improve the model.
- Model 6 is the final model for this branch in model development showing models all correlations between age, weight, height and BMI and posthoc eta for all parameters were $p>0.01$. The AIC was -6354.44 with an average MdAPE from cross-validation of 22.6%.
- Adding allometric scaling to the initial model (Model 7) improved the model.
- Adding aging to this model (Model 8) resulted in an improved model.
- Adding compartmental allometry (Model 9) to the allometric scaled model improved AIC but predictive performance in cross-validation was unchanged.
- Adding aging to this model (Model 10) resulted in an improved model.
- The addition of maturation functions based in age (Model 11) improved the model but basing maturation function on weight (Model 12) improved the model to a greater degree.
- Increasing CL for females (Model 13) improved AIC but decreased predictive performance in cross-validation. However, increasing V2, CL and Q2 for females aged 12-45 years (Model 14) improved the model to a greater degree.
- Model 14 show correlation between weight and in posthoc eta for V3. Adding an exponential weight correction function for V3 (Model 15) improved the model.
- Model 15 is the final model for this branch in model development showing models all correlations between age, weight, height and BMI and posthoc eta for all parameters were $p>0.01$. The AIC was -6512.78 with an average MdAPE from cross-validation of 20.9%.

The summarized findings of model development for weight scaling using and (estimated) BMI corrected weight are:

1) Including allometric scaling and compartmental allometry resulted in the best performing final model, both in terms of AIC and lowest MdAPE for cross-validation.
2) Maturation functions based on weight were only beneficial with allometric scaling and compartmental allometry.

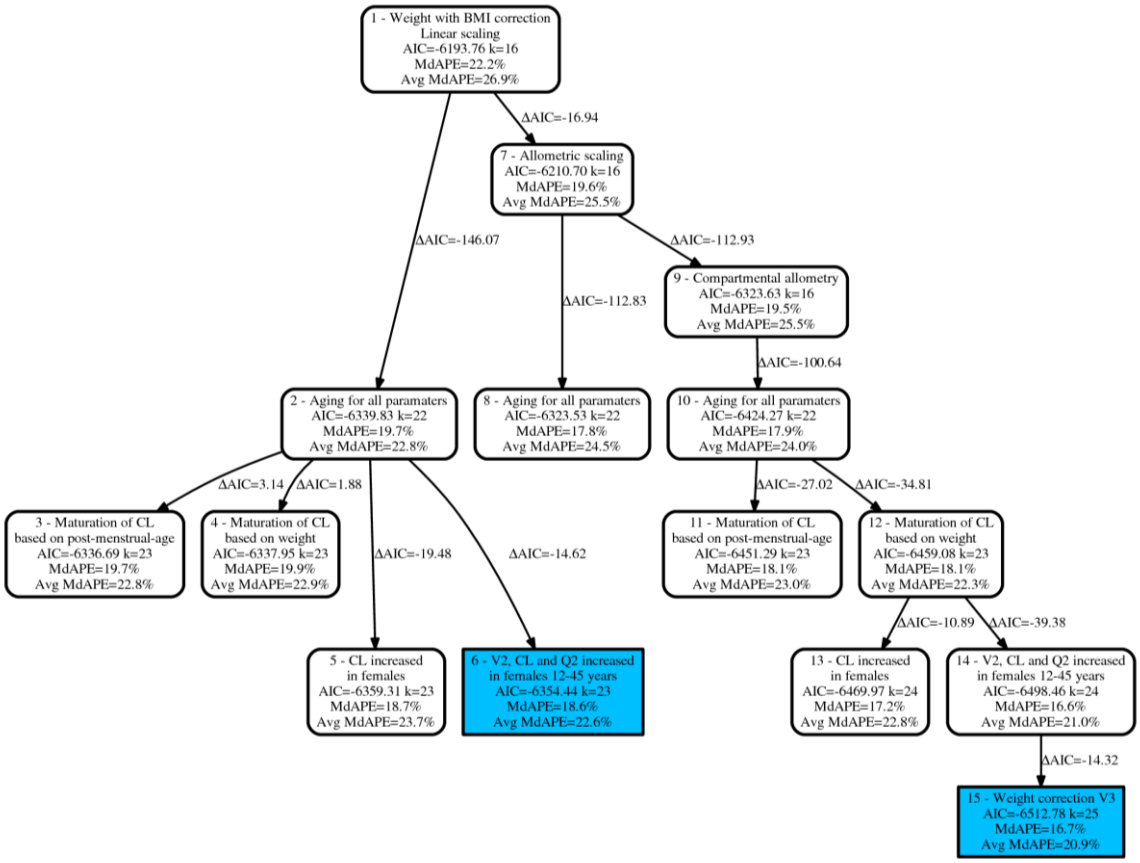

Figure 6. Hierarchical PK model building for size scaling using BMI-corrected weight.

General observations

1. Size scaling using the Al-sallami FFM equation with compartmental allometry leads to the best fitting model overall with the lowest objective function value and lowest average MdAPE over the subgroups in cross-validation. This is the size scaling function we used in the manuscript.
2. As expected the Janmahasatian FFM equation performs very similarly but slightly poorer compared to the Al-sallami FFM equation. In our opinion, there is no reason to use the Janmahasatian FFM equation because it is biased for small individuals and for adults it is identical to the Al-sallami FFM equation.
3. For most scaling methods (4 of 6) allometric scaling and compartmental allometry shows a clear improvement in AIC and prediction accuracy in cross-validation.
4. For most scaling methods (5 of 6) the best performing final model uses allometric scaling and compartmental allometry. In the 1 case where allometric scaling was not clearly the best model, it had a clearly better AIC (6373.91 vs. 6377.10) and only slightly poorer average MdAPE over the subgroups in cross-validation (21.8% vs. 21.5%).
5. For all scaling functions, compartmental allometry leads to a large improvement in AIC, around 100, compared to simple allometric scaling. This is true even in the cases where simple allometric scaling did not show a benefit.
6. For all scaling methods, the best performing final models use a maturation correction function on clearance.

7. In all cases where a maturation function was found beneficial, basing maturation on weight resulted in a better model than basing maturation on post-menstrual-age.
8. The scaling functions which incorporate sex (Al-sallami and Janmahastian) benefit the most from the sex correction step. Before this step these scaling functions perform slightly worse compared to sex-independent scaling methods. However, after the sex-correction the Al-sallami and Janmahastian FFM scaling function performed best.
9. The Dubois surface area scaling performs the best for the initial model tested. However model development "stalls" much earlier compared to other scaling functions. The final model for Dubois surface area scaling is the poorest of all scaling methods considered.
10. For linear weight scaling, a weight correction is needed for clearance both with and without allometric scaling. This suggests that linear weight scaling is not suitable for the data.
11. Only the Dubois surface area scaling function does not benefit from incorporating a weight correction to V3.

Conclusion

We believe that this comparison of model development for the size scaling functions justifies our use of size scaling using the Al-sallami FFM equation and allometric scaling since we found it to lead to the best performing final model.

Our previous experience with model development is repeated here: The model structure which provides the best initial steps in model development does not always lead to the best performing final model. Thus, a full model development process is necessary to rigorously evaluate a size scaling function.

---

[1] Al-Sallami HS, Goulding A, Grant A, Taylor R, Holford N, Duffull SB. Prediction of Fat-Free Mass in Children. Clin Pharmacokinet. 2015; 54(11): 1169-78.

2 Janmahasatian S, Duffull SB, Ash S, Ward LC, Byrne NM, Green B. Quantification of lean bodyweight. Clin Pharmacokinet, 2005; 44(10): 1051-65.

3 Du Bois D, Du Bois EF. A formula to estimate the approximate surface area if height and weight be known. 1916. Nutrition, 1989; 5(5): 303-11; discussion 312-3.