**Supplementary Material**

Accuracy of predicting chemical body composition of growing pigs using dual-energy X-ray absorptiometry

C. Kasper a, P. Schlegel a, I. Ruiz-Ascacibar a, 1 P. Stoll a, and G. Bee a

a Swine Research Unit, Agroscope, Tioleyre 4, 1725 Posieux, Switzerland

1 Present address: Yara Iberian S.A.U., Infanta Mercedes St., 28020 Madrid, Spain

Contents

[1. Raw data check of chemical analysis data 2](#_Toc50481544)

[2. Dietary treatments 2](#_Toc50481545)

[3. Model diagnostics 3](#_Toc50481546)

[4. Regressions on percentages 4](#_Toc50481547)

[5. Code for analyses in R 4](#_Toc50481548)

[6. References 5](#_Toc50481549)

# Raw data check of chemical analysis data

Chemical data were first checked by subtracting the sums of ash, crude protein, fat and water from the body weight obtained by scales. The mean difference in empty body weight was -57.44 ± 924.49 g and in carcass weight -63.91 ± 408.71 g. In addition, we visually examined the histograms of the proportions of water, ash, CP and lipid of the total sum of these components in connection with the coefficients of variation to check the distribution of the data and (Fig. S1).



**Figure S1:** Histograms and coefficients of variation (CV) of the proportions of water (red), ash (green), CP (yellow) and lipid (orange) determined by wet chemistry in the empty body (upper row, dark colours) and carcass (lower row, light colours) of entire male pigs. For code and data underlying the figure, see Kasper et al. (2020) and Supplementary Material section 6.

# Dietary treatments

Data used in this study originates from a so far unpublished experiment in which pigs were fed two grower and two finisher diets differing in their CP content (see Materials and Methods section in the main text for details). In this article, we were not specifically interested in the effect of dietary treatment. It was thus necessary to verify whether dietary treatments affected the estimates of slopes and intercepts of the prediction equations for chemical body composition based on DXA values. Table S1 shows the influence of dietary treatments on intercepts and slopes and the p-values of the treatment term. The treatment effect was only significant in the model of the energy content of the carcass, but the R2-value of the model including the treatment effect was only marginally higher than in the uncorrected model. Intercepts differed slightly between models containing the treatment effect and those that did not, but slopes remained largely unchanged (except for ash and lipid). Therefore, we concluded that, for the purpose of the present study, treatment could be omitted from the regression equations without compromising the estimates of slopes and intercepts.

**Table S1:** Effects of the inclusion of dietary treatment in the prediction equations on intercept, slope(s) and R2. Δintercept: difference in intercept, Δslope: difference in slope, ΔR2: difference in R2 of corrected model (including treatment effect) minus uncorrected model (without treatment effect). p-value: p-value of the treatment effect when included in the model.

|  |  |  |
| --- | --- | --- |
|  | **empty bodya** | **carcass** |
|  | Δ**Intercept** | Δ**Slope** | Δ**R2** | **p-value** | Δ**Intercept** | Δ**Slope** | Δ**R2** | **p-value** |
| weight (g) | 17.963 |  0.000 | 0.000 | 0.051 | 12.997 |  0.000 (mass) | 0.000 | 0.134 |
| energy content (MJ) | 0.318 |  0.000 (lean)-0.001 (fat) | 0.000 | 0.604 | 0.723 |  0.000 (lean) -0.001 (fat) | 0.001 | 0.039 |
| water (g) | -29.298 |  -0.001 (lean) | 0.000 | 0.257 | -41.839 |  -0.002 (lean) | 0.000 | 0.062 |
| ash (g) | -23.59 |  0.010 (BMC) | 0.000 | 0.387 | 3.438 |  -0.012 (BMC)0.001 (lean) | -0.002 | 0.793 |
| Ca (g) | 0.198 |  0.000 (BMC) | -0.003 | 0.985 | 2.622 | 0.001 (BMC) | -0.003 | 0.808 |
| P (g) | 0.911 |  -0.008 (BMC) 0.000 (lean) | -0.001 | 0.572 | 2.451 | - 0.009(BMC)0.001 (lean) | 0.000 | 0.299 |
| CP (g) | 2.821 |  0.000 (lean) | 0.000 | 0.890 | 3.775 |  0.000 (lean) | 0.000 | 0.849 |
| N (g) | 0.451 |  0.000 (lean) | 0.000 | 0.890 | 0.604 |  0.000 (lean) | 0.000 | 0.849 |
| lipid (g) | 9.145 | 0.011 (fat) -0.003 (lean) | 0.000 | 0.621 | 18.189 |  -0.030 (fat)0.011 (lean) | 0.001 | 0.091 |

aempty body contents at the time of live DXA scans estimated from empty body contents at slaughter

# Model diagnostics

Model diagnostics were carried out using the packages car (version 3.0-8; Fox & Weisberg, 2019) and MASS (version 7.3-51.6; Venables & Ripley, 2002) in R (version 3.6.3; R Core Team, 2020). They revealed that some assumptions of linear models were violated. In particular, the assumption of homogeneity of variance was not met judging from visual inspection of residuals vs. fitted values. Variances for larger fitted values were higher than those of lower fitted values in all models. In some of the models[[1]](#footnote-1), we detected statistical outliers, but the assessment of the Cook's distance showed that none of them seemed to have a pronounced influence on the models. Removal of those outliers never resulted in improved model diagnostics. One individual (8729-FE3), which was an outlier in five of the models (ash, Ca and P), was ranked second highest in ash, Ca and P content per g body weight. Another individual (9122-FE3), which was an outlier in the model predicting water from lean mass, had a rather high, but not extreme, lean mass in DXA. Both individuals did not have unusual chemical or DXA values and there was no indication of errors, thus we did not remove these individuals from the dataset. Visual checks revealed that residuals were more or less normally distributed. The lm() procedure in R is considered quite robust, but fitting models with linear least-squares regression in order to derive the parameters of the regression line while violating core assumptions might bear the risk of obtaining unreliable slopes and intercepts for our prediction equations. Thus, we used a "leave-one-out" cross-validation procedure in the package caret (version 6.0-86; Kuhn, 2020) to construct a non-parametric confidence interval for the slopes and intercepts, which are presented in Table 3 in the main text. To obtain the confidence intervals, we computed intercepts and slopes of all models by omitting one row at a time, ordered the estimates according to their value and took the 3rd and 59th and the 3rd and 66th ranked estimates for models of the empty body and the carcass, respectively, as the limits of the confidence intervals. This corresponds to a 93% CI in the empty body and a 94% CI in the carcass.

# Regressions on percentages

For a comparison with the percentages of lean tissue mass and lipid in the carcass halves by (Mitchell et al., 2003), we built separate prediction equations. We regressed the percentage of lean tissue by DXA on the percentage of lean tissue mass, which was calculated as the sum of CP and water divided by sum of CP, lipid, water and ash of the half carcasses. The percentage of fat tissue mass by DXA was regressed on the percentage of lipid content in the half carcasses (lipid divided by sum of CP, lipid, water and ash). Table S2 shows the prediction equations derived for % lean tissueDXA and % lipidDXA.

**Table S2:** Summary of model estimates of percentages in the half-carcasses including their standard error, their p-value, R2 and RMSE from back-prediction of chemical values.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **chemical variable (predicted variable)** | **term (DXA variable)** | **estimate (SE)** | **p-value** | **R2** | **RMSE** |
| **% lean tissue (CP + water)****(% LeanDXA)** | Intercept % lean | -7.414 (4.501)1.061 (0.056) | 0.104< 0.001 | 0.842 | 1.40 |
| **% lipid****(% lipidDXA)** | Intercept % fat  | 0.565 (0.957)1.074 (0.057) | 0.557< 0.001 | 0.841 | 1.41 |

# Code for analyses in R

*[see Kasper\_etal\_DXA.Rmd as a separate file]*

# References

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (Third). Sage. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>

Kasper, C., Schlegel, P., Ruiz-Ascacibar, I., Stoll, P. and Bee, G., 2020. Data of "Accuracy of predicting chemical body composition of growing pigs by dual-energy X-ray absorptiometry". Zenodo. DOI: 10.5281/zenodo.3981182.

Kuhn, M. (2020). *caret: Classification and Regression Training*. https://CRAN.R-project.org/package=caret

Mitchell, A. D., Scholz, A. M., & Pursel, V. G. (2003). Prediction of pork carcass composition based on cross-sectional region analysis of dual energy X-ray absorptiometry (DXA) scans. *Meat Science*, *63*(2), 265–271. https://doi.org/10.1016/S0309-1740(02)00081-5

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (Fourth). Springer. http://www.stats.ox.ac.uk/pub/MASS4

1. Body weight and Ca in the carcass, energy, water, ash content in empty body, Ca, P in empty body; empty body ash and lipid contents predicted from carcass [↑](#footnote-ref-1)