## modeling in physiology

# Scaling, normalizing, and per ratio standards: an allometric modeling approach

ALAN M. NEVILL AND ROGER L. HOLDER

School of Sport and Exercise Sciences and School of Mathematics and Statistics, University of Birmingham, Birmingham B15 2TT, United Kingdom

Nevill, Alan M., and Roger L. Holder. Scaling, normalizing, and per ratio standards: an allometric modeling approach. J. Appl. Physiol. 79(3): 1027-1031, 1995.—The practice of scaling or normalizing physiological variables (Y) by dividing the variable by an appropriate body size variable (X) to produce what is known as a "per ratio standard" (Y/ X), has come under strong criticism from various authors. These authors propose an alternative regression standard based on the linear regression of (Y) on (X) as the predictor variable. However, if linear regression is to be used to adjust such physiological measurements (Y), the residual errors should have a constant variance and, in order to carry out parametric tests of significance, be normally distributed. Unfortunately, since neither of these assumptions appear to be satisfied for many physiological variables, e.g., maximum oxygen uptake, peak and mean power, an alternative approach is proposed of using allometric modeling where the concept of a ratio is an integral part of the model form. These allometric models naturally help to overcome the heteroscedasticity and skewness observed with per ratio variables. Furthermore, if per ratio standards are to be incorporated in regression models to predict other dependent variables, the allometric or log-linear model form is shown to be more appropriate than linear models. By using multiple regression, simply by taking logarithms of the dependent variable and entering the logarithmic transformed per ratio variables as separate independent variables, the resulting estimated log-linear multipleregression model will automatically provide the most appropriate per ratio standard to reflect the dependent variable, based on the proposed allometric model.

allometry; log-linear models; heteroscedasticity; multiple regression

TANNER (27), followed by various other authors (12–15), including most recently Toth et al. (28), have written about the problems associated with using a particular type of ratio, sometimes referred to as "per ratio standards," when scaling or normalizing the results of selected measurements in physiology and clinical medicine. In these disciplines, it is common practice to express various measurements (Y), such as oxygen consumption  $(\mathring{V}O_2)$  and cardiac output, as per-weight or per-surface area ratios, since by dividing by an appropriate body-size variable (X) it is assumed that differences due to the subject's body size will have been removed, i.e., the ratio standard (Y/X) is assumed to have scaled or normalized the variable (Y) to be independent of the body-size variable (X).

In addition, Tanner (27) and supporters (12-14) ob-

served that many of these per ratio standards fail to render the measurements independent of body size, e.g., when maximum oxygen uptake  $(\dot{V}_{O_{2\,max}}; l/min)$  is recorded per body weight  $(ml \cdot kg^{-1} \cdot min^{-1})$  and then correlated with body weight, a significant negative correlation is invariably found. Furthermore, these authors argue that, if the ratio standard (Y/X) were a true linear proportion, then when the physiological variable (Y) was regressed against the predictor body-size variable (X), the resulting linear regression equation should have an intercept statistic close to zero. Because this assumption is rarely satisfied. Tanner (27), and later Katch (12) proposed an alternative regression standard to represent the subject's body size-adjusted measurement. These regression standards are obtained by adding to the group mean the subject's residual error, taken from the regression line. Clearly, the assumptions associated with the use of regression standards as proposed by Tanner and supporters are 1) that the relationship between (Y) and (X) is linear, i.e., given by

$$Y = a + b \cdot X + \epsilon \tag{1}$$

and 2) that the error term  $\epsilon$  has constant variance throughout the range of observations.

Toth et al.(28) also advocate the use of regression standards when normalizing or adjusting peak  $\dot{V}O_2$  data (Y) for differences in fat-free mass (FFM; X). The variable FFM was chosen in preference to other body composition variables (e.g., body mass or fat mass), since it was found to be the best single predictor of peak  $\dot{V}O_2$ . When the linear regression model (Eq. 1) was fitted separately to the 322 males (aged 17–78 yr) and 201 females (aged 18–81 yr), the intercept statistics for both males (-1.018  $\pm$  0.33) and females (-1.078  $\pm$  0.31) were significantly different from zero. The authors argue that the existence of significant intercept parameters justifies the use of regression standards in preference to ratio standards.

In a recent article, Kronmal (15) restates Tanner's concerns with the use of ratio standards. In addition, he draws attention to the work of Pearson (24) and, later, Neyman (23), who originally observed that spuriously high correlations are found between indexes that have a common component. The author goes on to recommend that a ratio standard (Y/X) should only be incorporated into regression analyses as part of a full linear model when the constituent parts that make up the ratio, i.e., (Y) and  $(X^{-1})$ , are also included as main

effects, stating "it is not good practice to include interactions in an equation without first including the variables that comprise it as first-order terms in the model." As with the regression standard models proposed by Tanner (27) and Katch (13, 14), Kronmal (15) assumes that the models discussed in his article are linear with additive components and have an error term that has constant variance throughout the range of observations.

When using multilevel modeling procedures to investigate the development of aerobic power in young athletes, Baxter-Jones et al. (3) also assume that  $\dot{V}O_{2\,max}$  has an additive model and error structure. However, since  $\dot{V}O_{2\,max}$  is known to be proportional to body size, an additive model is unlikely to satisfactorily normalize  $\dot{V}O_{2\,max}$  for differences in the chosen body-size variables. By observing their Table 3, the groups' mean scores for  $\dot{V}O_{2\,max}$ , whether recorded in the units of liters per minute or milliliters per kilogram per minute, appear to be related and proportional to their SD values.

### ALLOMETRIC MODELING

Caution should be exercised when using linear regression models to investigate the relationship between a ratio standard (Y/X) as the dependent variable and then incorporating a body-size variable (X) as an independent variable. These regression models will tend to produce biased results due to correlated random variation in the dependent and independent variables and this is, in essence, the classic phenomenon of regression to the mean. Another possible cause for concern when using linear regression to model per ratio variables is that the variance of the error term may not be constant throughout the range of observations. This concern appears to be justified when observing two of the examples discussed above. In Fig. 1 in Toth et al. (28) and Table 3 in Baxter-Jones et al. (3), there is clear evidence of heteroscedasticity that contradicts an assumption of a constant error variance, i.e., in both these examples, the error variation appears to increase in proportion to

However, when considering such problems associated with scaling and per ratio standards, a number of authors (e.g., see Refs. 5, 6, 9, 16, 17, 22, 25) have recognized the value of an important class of models, often referred to as "allometric or power function models." For these models, the concept of a ratio is an integral part of the model form, and the variables and errors are assumed to be proportional and multiplicative, respectively. Consequently, when a logarithmic transformation is applied, the model becomes a "log-linear regression model." In general, a log-linear (regression) model is a regression model in which the dependent variable is log-transformed, and the independent variable terms are in a form that is linear in the unknown parameters. In its untransformed form, the model will be the product of the independent variable terms (e.g., power and exponential functions) and the error term. Standard regression methodology may be used to estimate the unknown parameters, provided the log of the error term is independent and has a normal distribution with constant variance.

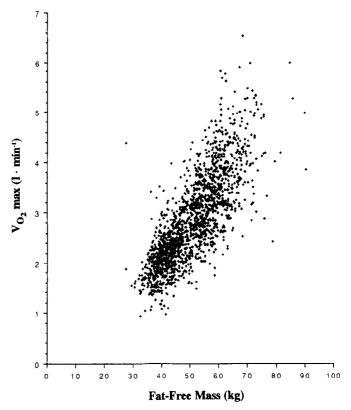


FIG. 1. Maximum oxygen uptake  $(\dot{V}O_{2\,max})$  vs. fat-free mass for 1,734 subjects from the Allied Dunbar National Fitness Survey.

Above all other reasons, a multiplicative model has a theoretical aptness in many scale- and size-related contexts. Indeed, as the dependent variable often covers a theoretically unbounded order of magnitude and is, therefore, likely to produce a positively skewed distribution, the logarithmic transformation will naturally provide both a linear model and a more symetrically distributed dependent variable.

With such a large age range, the heteroscedasticity observed in Fig. 1 of Toth et al. (28) may have been partially caused by omitting the variable age. Indeed, when using the results from the Allied Dunbar National Fitness Survey (ADNFS; Ref. 26), Nevill and Holder (20) discovered that both body size and age were essential components in the  $\dot{V}O_{2\,max}$  regression models. When  $\dot{V}O_{2\,max}$  was plotted against estimated FFM for the ADNFS data (Fig. 1), a similar pattern was observed to that of Toth et al. (28), suggesting that the two data sets are comparable. Estimated FFM was calculated from percentage body fat, the latter being derived by using the methods of Durnin and Womersley (8).

## ALLOMETRIC MODELS WITH PER RATIO STANDARD VARIABLES AS THE DEPENDENT VARIABLES

Various authors have reported heteroscedasticity in their data (e.g., see Refs. 11, 20-22) and have proposed the allometric (or power function) model

$$Y = a \cdot X^b \cdot \epsilon \tag{2}$$

to explain such relationships. Not only does the model explain the proportionality between the variables Y

and X but the multiplicative error term is better able to describe the heteroscedasticity associated with such data. Support for the log-transformation of model (Eq. 2) can be obtained by using transformation methods of Box and Cox (4). In the cases studied by Nevill and Holder (20, 21), when investigating the relationship between  $\dot{V}O_{2\,\text{max}}$ , body weight, lean body mass, and height, the logarithmic transformation was found to be the most successful at overcoming heteroscedasticity and providing normally distributed errors.

When scaling or normalizing various physiological measurements for individuals of different body size, Nevill et al. (22) proposed the power function ratio standard  $(Y/X^b)$ , derived from the allometric model (Eq. 2) that would render the subject's physiological performance variable (Y) independent of their body-size variable (X). The authors were able to confirm empirically [anticipated by Astrand and Rodahl (1)] that  $\dot{V}O_{2\,\text{max}}$  was proportional to body mass,  $m^{2/3}$ , and hence should be scaled by recording  $\dot{V}O_{2\,\text{max}}$  in the units of  $(\text{ml} \cdot \text{kg}^{-2/3} \cdot \text{min}^{-1})$  to be independent of body size. They were also able to demonstrate that both peak and mean power output (W) needed to be scaled by recording peak and mean power measured in  $(W \cdot \text{kg}^{-2/3})$  to be independent of body size.

Simply by taking logarithms of the power function ratio standard  $(Y/X^b)$ , Nevill et al. (22) were able to show that further experimental design effects can be investigated using traditional analysis of variance methods. The same design effects can be obtained using the analysis of covariance, by analyzing  $\log(Y)$  as the dependent variable, where the covariate  $\log(X)$  would automatically adjust the numerator dependent variable to allow for their association. Further covariates can be easily incorporated in the allometric model, for example Nevill and Holder (20) incorporated the variable "age" as an exponential term that had the additional operational advantages of providing a plausible asymptotic age decline in  $VO_{2 \max}$  and, after a logarithmic transformation, a linear model structure. The proposed model for  $VO_{2 \max}$  was given by

$$\dot{\mathbf{V}}_{O_{2\,\mathrm{max}}} = \mathrm{weight}^k \cdot \exp(c + d \cdot \mathrm{age}) \cdot \epsilon \tag{3}$$

that can be linearized with a log-transformation to produce the following log-linear model

$$\log_{e} (\dot{\mathbf{V}}_{O_{2 \max}})$$

$$= k \cdot \log_{e} (\text{weight}) + c + d \cdot \text{age} + \log_{e} (\epsilon)$$
(4)

When the log-linear model (Eq.~4) was fitted to the results of the ADNFS (26), by using standard linear regression methodology as implemented in MINITAB (18), the resulting solution explained  $R^2=74.1\%$  of the variation in  $\dot{V}o_{2\,max}$ . The fitted model retained just the single-exponent parameter for weight k=0.66 (SE = 0.025) agreeing with the anticipated  $k={}^2/_3$ , based on theoretical and physiological grounds. The residual errors from fitting the log-linear model (Eq.~4) were also found to be acceptably normal when using the probability plot correlation test for normality (10), also available in MINITAB (18).

As mentioned above, the allometric model structure

can be easily extended to include further categorical and quantitative terms. For example, model in Eq. 3 can be developed to incorporate FFM in place of body weight and to include an additional age<sup>2</sup> term, as follows

$$\dot{\mathbf{V}}_{\mathbf{O}_{2\,\text{max}}} = \mathbf{F}\mathbf{F}\mathbf{M}^k \cdot \exp(c + d \cdot \mathbf{age} + e \cdot \mathbf{age}^2) \cdot \epsilon \quad (5)$$

After taking the log-transformation of model (Eq. 5), the log-linear model was then fitted to the results of the ADNFS (26) by using the methods of Nevill (19) that required a "developmental" component for body size. The resulting solution explained  $R^2 = 75.3\%$  of the variation in  $VO_{2 max}$ . As before, the exponent for FFM was k = 0.66 (SE = 0.056), supporting the anticipated parameter  $k = \sqrt[2]{3}$ , based on theoretical and physiological grounds. The age2 term made a significant contribution to the model, and a subsequent examination of the fitted parameters indicated that  $\dot{V}_{O_{2\,max}}$  reached a peak in the subjects' early 20s and subsequently declined thereafter. Once again, the residual errors from the log-linear model (Eq. 5) were found to be acceptably normal when using the probability plot correlation test for normality.

When model in Eq. 5 was fitted to the ADNFS (26) data, omitting the age and  $age^2$  terms, the resulting solution explained  $R^2=68.4\%$  of the variation in  $\dot{V}_{O_{2\,max}}$ , with the exponent for the term FFM increasing implausibly to k=1.06 (SE = 0.060). Indeed, when the model proposed by Toth et al. (28) was fitted to the male and female ADNFS subjects separately, by using the linear rather than allometric model, the resulting solution explained just  $R^2=65.0\%$  of the variation in  $\dot{V}_{O_{2\,max}}$ . Furthermore, the residuals from fitting this model were found to be nonnormal when the probability plot correlation test for normality was used, as described previously.

## ALLOMETRIC MODELS INCORPORATING PER RATIO STANDARD VARIABLES AS THE INDEPENDENT VARIABLES

Because per ratio standard variables are known to be proportionally related, Nevill et al. (22) used the multiplicative allometric model (Eq. 6) to investigate the relationship between 5-km run speed (Z),  $\dot{V}_{O_{2\,max}}(Y)$ , and body weight (X) of 204 recreationally active subjects (n=112 men; n=92 women)

$$Z(m/s) = a \cdot Y^{b_1} \cdot X^{b_2} \cdot \epsilon \tag{6}$$

Support for the use of the multiplicative (allometric) model (Eq.~6) was obtained when running speed was plotted against  $\dot{V}_{O_{2\,max}}$  (Fig. 2), when heteroscedasticity was observed clearly for both the male and female subjects.

When we used Generalized Linear Interactive Modeling (2) using multiple log-linear regression, we found no statistically significant differences between the male and female models for both the  $\dot{\text{Vo}}_{2\,\text{max}}$  and the body mass a and b parameters. Hence, the common power function model relating 5-km running speed, Z (m/s), to  $\dot{\text{Vo}}_{2\,\text{max}}$  Y (l/min) and body mass X (kg) is given by

$$Z \text{ (m/s)} = 84.3 \cdot Y^{1.01} \cdot X^{-1.03} \tag{7}$$

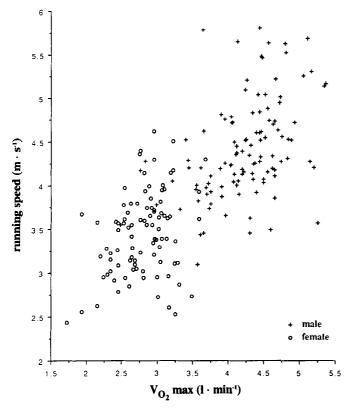


FIG. 2. Five-kilometer run speed (m/s) vs.  $\dot{V}O_{2 \text{ max}}$  for 204 recreationally active subjects (men: n=112; women: n=92).

The parsimonious solution given by Eq.~7 is both simple and meaningful. The fitted predictor of 5-km run times, when recorded as a rate of performance, i.e., mean running speed (m/s), is almost exactly proportional to the ratio standard  $\dot{V}_{O_{2\,max}}$  (l/min) divided by body mass (kg) or the well-known weight-related  $\dot{V}_{O_{2\,max}}$  (ml·kg<sup>-1</sup>·min<sup>-1</sup>). The alternative full linear model proposed by Kronmal (15) to describe the running speed results Z would be

$$Z = a + b_1 \cdot Y + b_2 \cdot X^{-1} + b_3 \cdot Y \cdot X^{-1} + \epsilon$$
 (8)

To compare the allometric log linear model (Eq. 6) with the equivalent full linear model (Eq. 8), the criterion of Cox (7) based on the difference between the two models' maximized log-likelihoods was used. The maximized log-likelihood statistic for the allometric model (Eq. 6) was found to be -9.75 using just three parameters. In contrast, the maximized log-likelihood statistic for the full linear model (Eq. 8) was less at -11.21 requiring four parameters, suggesting the superiority of the allometric model (Eq. 6) over the linear model (Eq. 8). However, it is also important to note that, although model in Eq. 8 fits the data moderately well, it has no obvious physiological interpretation, unlike the allometric model (Eq. 6).

### DISCUSSION

The use of per ratio standards has come under strong criticism from a number of authors, most recently in articles by Toth et al. (28). These authors argue that rather than using a per ratio standard as the depen-

dent variable in a regression analysis, the numerator of the per ratio standard alone should be defined as the dependent variable, allowing the denominator (body size) variable to be incorporated as a covariate (independent variable) in the regression model. From the fitted regression model, a regression standard can be obtained

If linear regression methods are to be adopted when scaling or normalizing physiological variables, an important assumption associated with such methods is that the residual errors should have a constant variance. Furthermore, in order to carry out parametric tests of significance, it is necessary that these errors should have a normal distribution. Unfortunately, neither of these requirements appear to be satisfied for body-size-related variables such as  $\dot{V}O_{2\,max}$ . It would appear that allometric models are more appropriate for investigating such body-size-related variables. As well as their theoretical aptness, these types of models naturally overcome the heteroscedasticity associated with per ratio variables and produce normally distributed residual errors.

More generally, if per ratio standards are to be incorporated in regression models as independent variables, the allometric model form, e.g., Eq. 6, should be considered. This can be achieved simply by taking logarithms of the performance-dependent variable and entering the logarithmic transformed physiological and body-size variables as separate independent variables in a multiple log-linear regression. The logarithmic transformation appears to overcome the problems associated with linear regression, i.e., heteroscedasticity and non-normality, and the resulting log-linear multiple regression model will automatically provide the most appropriate fitted per ratio standard to reflect the dependent variable.

The authors thank The Sports Council and the Health Education Authority (United Kingdom) for access and permission to publish the results from the Allied Dunbar National Fitness Survey. We are also grateful to the editor and the three referees for their valuable suggestions in revising the manuscript.

Address for reprint requests: School of Sport and Exercise Sciences, Univ. of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom.

Received 25 July 1994; accepted in final form 4 May 1995.

## REFERENCES

- Astrand, P.-O., and K. Rodahl. Textbook of Work Physiology (3rd ed.). New York: McGraw-Hill, 1986.
- Baker, R. J., and J. A. Nelder. Generalized Linear Interactive Modelling, Release 3. Oxford, UK: Numerical Algorithms Group, 1978
- Baxter-Jones, A., H. Goldstein, and P. Helms. The development of aerobic power in young athletes. J. Appl. Physiol. 75: 1160-1167, 1993.
- 4. Box, G. E. P., and D. R. Cox. An analysis of transformations (with Discussion). J. R. Statist. Soc. B 26: 211-252, 1964.
- Butler, J. P., H. A. Feldman, and J. J. Fredberg. Dimension analysis does not determine a mass exponent for metabolic scaling. Am. J. Physiol. 245 (Regulatory Integrative Comp. Physiol. 14): R195-R199, 1987.
- Calder, W. A., III. Size, Function, and Life History. Cambridge, MA: Harvard Univ. Press, 1984.

- Cox, D. R. Further results on tests of separate families of hyotheses. J. R. Statist. Soc. B 24: 406-424, 1962.
- 8. **Durnin, J. V., and J. Womersley.** Body fat assessed from total body density and its estimation from skinfold thickness measurements on 481 men and women aged 16 to 72 years. *Br. J. Nutr.* 32: 77–85, 1974.
- Feldman, H. A. On the allometric mass exponent, when it exists. J. Theor. Biol. 172: 187-197, 1995.
- Filliben, J. J. The probability plot correlation coefficient test for normality. *Technometrics* 17: 111-117, 1975.
- Jolicoeur, H., and A. A. Heusner. The allometry equation in the analysis of the standard oxygen consumption and body weight of the white rat. *Biometrics* 27: 841-855, 1971.
- 12. **Katch, V. L.** Correlation v ratio adjustment of body weight in exercise-oxygen studies. *Ergonomics* 15: 671-680, 1972.
- Katch, V. L. Use of the oxygen/body weight ratio in correlational analyses: spurious correlations and statistical considerations. Med. Sci. Sports Exercise 5: 253-257, 1973.
- 14. Katch, V. L., and F. I. Katch. Use of weight-adjusted oxygen uptake scores that avoid spurious correlations. Res. Q. 4: 447– 451, 1974.
- Kronmal, R. A. Spurious correlation, and the fallacy of ratio standard revisited. J. R. Statist. Soc. A 156: 379-392, 1993.
- McMahon, T. A. Muscles, Reflexes, and Locomotion. Princeton, NJ: Princeton Univ. Press, 1984.
- McMahon, T. A, and J. T. Bonner. On Size and Life. New York: Scientific American Books, 1983.
- 18. MINITAB, Inc. MINITAB Reference Manual. 1989.

- Nevill, A. M. The need to scale for differences in body size and mass: an explanation of Kleiber's 0.75 mass exponent. J. Appl. Physiol. 77: 2870-2873, 1994.
- Nevill, A. M., and R. L. Holder. Modelling maximum oxygen uptake: a case study in non-linear regression formulation and comparison. Appl. Statist. 43: 653-666, 1994.
- 21. Nevill, A. M., and R. L. Holder. Body mass index: a measure of fatness or leanness? Br. J. Nutr. 73: 507-516.
- Nevill, A. M., R. Ramsbottom, and C. Williams. Scaling physiological measurements for individuals of different body size. Eur. J. Appl. Physiol. Occup. Physiol. 65: 110-117, 1992.
- 23. Neyman, J. Lectures and Conferences on Mathematical Statistics and Probability (2nd ed.). Washington, DC: US Department of Agriculture, 1952, p. 143-154.
- 24. **Pearson, K.** Mathematical contributions to the theory of evolution-on a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proc. R. Soc. Lond.* 60: 489-497, 1897.
- 25. Schmidt-Nielsen, K. Scaling: Why Is Animal Size so Important? Cambridge, UK: Cambridge Univ. Press, 1984.
- Sports Council. Allied Dunbar National Fitness Survey: Main Findings. London: Sports Council, 1992.
- Tanner, J. M. Fallacy of per-weight and per-surface area standards and their relation to spurious correlations. J. Appl. Physiol. 2: 1-15, 1949.
- 28. Toth, M. J., M. I. Goran, P. A. Ades, D. B. Howard, and E. T. Poehlman. Examination of data normalization procedures for expressing peak Vo<sub>2</sub> data. J. Appl. Physiol. 75: 2288-2292, 1993.

