

The paradigm of body condition: a critical reappraisal of current methods based on mass and length

Jordi Peig^{1,*} and Andy J. Green²

¹Department of Animal Biology, University of Barcelona, Avinguda Diagonal 08028 Barcelona, Spain; and ²Department of Wetland Ecology, Estación Biológica de Doñana – CSIC, Américo Vespucio, 41092 Sevilla, Spain

Summary

1. Body condition is a major concept in ecology addressed in countless studies, and a variety of non-destructive methods are used to estimate the condition of individuals based on the relationship between body mass M and measures of length L . There is currently no consensus about the most appropriate condition index (CI) method, and various traditions have been established within subdisciplines in which ecologists tend to apply that method used previously by their peers.

2. Here, we present a reappraisal of six conventional CI methods: Fulton's index ($K = M/L^3$), Quételet's index ($BMI = M/L^2$), Relative condition (K_n , computed as the observed individual mass divided by the predicted mass $M_i^* = a L_i^b$ where a and b are determined by ordinary least squares (OLS) regression of M against L), Relative mass (W_r , where a and b above are determined from a reference population), the Residual index (R_i , the residuals from an OLS regression of M against L) and ANCOVA. We compare the performance of these methods with that of the Scaled mass index, a novel method which was previously shown to perform better than R_i as a predictor of fat and other body components [J. Peig & A.J. Green (2009) *Oikos*, **118**, 1883].

3. To be reliable, a CI method must successfully account for the changing relationship between M and L as body size changes and growth occurs (i.e. for the scaling relationship between M and L). Using data from three species of small mammals we show that, unlike the Scaled mass index, all six conventional methods fail to do this, and as a result they consistently lead to significant differences in CIs between age classes and sex that are a mere consequence of changes in body size. The Scaled mass index was also particularly successful at detecting changes in CI resulting from high levels of contaminants.

Key-words: allometry, body size, condition index, fitness, Scaled mass index, standardized major axis

Introduction

Body condition is intimately related to an animal's health, quality or vigour (Peig & Green 2009), and has been widely claimed to be an important determinant of fitness. A wide range of morphological, biochemical or physiological metrics have been proposed as condition indices (CIs) (Stevenson & Woods 2006). Here, we are only concerned with CIs based on the relationship between body mass (M) and length measurements (L), whose ultimate goal is to interpret variations of body mass for a given body size as an attribute of the individual's well-being (most typically, variation in the size of energy reserves).

A variety of formulas and statistical methods have been proposed to standardize body size, and there is much debate

about which ones are most suitable as CIs (Stevenson & Woods 2006). Among conventional methods, simple ratios (or ratio indices) between M and L or L raised to a specific power (e.g. M/L , M/L^2 , M/L^3) have been in use the longest. Examples include Fulton's index ' K ' (where $K = M/L^3$) still used in some ecological studies, or Quételet's index or the body mass index ($BMI = M/L^2$) universally applied in health sciences. In fisheries, a popular CI is the so-called Relative condition ' K_n ', computed as the observed individual mass (M_i) divided by the predicted mass (M_i^* , where $M_i^* = a L_i^b$). The estimates a and b are empirically determined by ordinary least squares (OLS) regression of M against L (both log-transformed) for the whole study population (LeCren 1951). Even more popular in fisheries is a variant of the K_n index, called Relative mass ' W_r ', where a and b are determined from a reference population instead of the population under study (Murphy, Willis & Springer 1991).

*Correspondence author. E-mail: jpeig@ub.edu

In recent years, the most widely accepted CI in terrestrial ecology has been the Residual index ' R_i ', which uses the residuals from an OLS regression of M against one or more length measurements, usually after log transformation (Jakob, Marshall & Uetz 1996; Schulte-Hostedde, Millar & Hickling 2001; Ardia 2005; Schulte-Hostedde *et al.* 2005). Another popular approach is to conduct an analysis of covariance (ANCOVA), which combines features of linear regression and ANOVA to estimate directly the treatment effect on M while controlling for a concomitant variable of influence, denoted by L (García-Berthou 2001; Velando & Alonso-Alvarez 2003; Serrano *et al.* 2008).

Currently, no consensus exists about the best CI or criteria which allow selection of the most appropriate method in a particular study, and few authors provide a detailed justification of their choice of method. Ecologists and epidemiologists follow traditions within their discipline. Hence, the continuing use of ratios in fisheries and health sciences or the widespread use of R_i in terrestrial ecology, despite criticisms of these approaches (Albrecht, Gelvin & Hartman 1993; Packard & Boardman 1999; García-Berthou 2001; Green 2001; Freckleton 2002). As we will demonstrate, results may differ dramatically depending on the method of choice, causing concern about the reliability of studies based on less appropriate methods.

Peig & Green (2009) presented a novel CI method called the Scaled mass index ' \hat{M}_i ', which standardizes body mass at a fixed value of a linear body measurement based on the scaling relationship between mass and length, according to equation 1:

Scaled mass index (\hat{M}_i):

$$\hat{M}_i = M_i \left[\frac{L_0}{L_i} \right]^{b_{\text{SMA}}} \quad \text{eqn 1}$$

where M_i and L_i are the body mass and linear body measurement of individual i respectively; b_{SMA} is the scaling exponent estimated by the standardized major axis (SMA) regression of $\ln M$ on $\ln L$; L_0 is an arbitrary value of L (e.g. the arithmetic mean value for the study population); and \hat{M}_i is the predicted body mass for individual i when the linear body measure is standardized to L_0 . Making a comparison of M_i values between different populations or studies simply requires use of the same L_0 value in eqn. 1. In a variety of vertebrate species (five small mammals, one bird and one snake), the Scaled mass index performed better than the Residual index ' R_i ' as a predictor of variations in fat and protein reserves as well as other body components (Peig & Green 2009).

The overall objective of this paper is to critically reappraise current CI methods and compare their performance with that of the Scaled mass index ' \hat{M}_i '. Using field data from small mammals, we compare the performance of \hat{M}_i , Residual index ' R_i ', ANCOVA, Body mass index 'BMI', Fulton's index ' K ', Relative condition ' K_n ', and Relative mass ' W_r '. We demonstrate empirically that conventional CIs can be inherently biased with regard to animal size, and tend to change condition scores in larger animals owing to violations of statistical

assumptions and failure to account for growth and scaling relationships.

We included ratio methods in our study, because they are still widely in use and have not previously been compared with \hat{M}_i . Although ratio methods have been widely criticized, many of their problems are shared by R_i and ANCOVA. Exploring the performance of ANCOVA was also important because, like \hat{M}_i , it has been advocated as a more reliable method than ' R_i ' (García-Berthou 2001). The ANCOVA is not strictly a CI but rather an inferential test where individual scores are absent, making validation via correlations with body components such as fat reserves impossible.

First, to help explain why CIs can be unreliable, we consider basic principles underlying the construction of condition indices based on mass and length data.

FIRST PRINCIPLES FOR A CONDITION INDEX

As pointed out by Kotiaho (1999), if CIs are assumed to reflect, or be validated by, either the 'absolute' or 'percentage' amount of energetic tissue, they should not be used for comparisons of condition among individuals of different size. This is because, at each stage of growth and development, there is both an optimal range of structural and energy capital (in absolute or % values) and an optimal distribution of this capital among different body components. In other words, the proportional or absolute amount of energy stores can be expected to change with normal growth processes, even in an optimal environment (e.g. one free from pathogens and disturbance, with non-limiting resources). It would therefore be erroneous to assert that adults were in better 'condition' or 'well being' than juveniles because of a greater absolute amount of fat or protein, or that juveniles were in better condition than adults because of a greater relative content (as %fat, %protein, etc.). The same generally applies to comparisons between sex or subspecies (Gallagher *et al.* 1996). Kotiaho's premise has profound implications for condition estimates and the search for a suitable method. Essentially, a method is required that accounts for normal growth processes (i.e. scaling), thus allowing a valid comparison between individuals of a different body size.

Growth leads to strong correlations between M and most linear body measurements (L) (Hoppeler & Weibel 2005; Peig & Green 2009). Because increases in both M and L are parts of the same growth phenomenon (Thompson 1961), both M and L are indicators of body size *per se* which require some type of mutual standardisation. Strictly speaking, however, as body growth involves not only variation in body size but also in body composition (Huxley 1932; Calder 1984), the aim of any CI is not to control for body size but rather for growth effects as a whole and their consequences for scaling (Peig & Green 2009).

Virtually all work on biological scaling assumes a power function of the form $Y = \alpha X^\beta$ where the parameter α is a constant and the parameter β , the 'allometric or scaling exponent', determines the dimensional balance between Y

and X . The power (nonlinear) relationship is supported by the notion of growth as a multiplicative process of living matter (Shea 1985). β equals the ratio of the specific growth rates of the dimensions Y and X (i.e. dY/Ydt and dX/Xdt , Shea 1985). As the same time interval applies for changes in M and L , estimating body condition from M - L data is time-independent and unaffected by the rate of growth.

Morphogenesis is the mechanism controlling the body proportions (i.e. between M and L) and thus determines the scaling exponent β (Roth & Mercer 2000). This process is largely controlled by genes which regulate how body parts differentiate and orientate to form a well-structured organism (Conlon & Raff 1999; Hogan 1999). As monomorphic species have their own body plan distinct from other species, the assumption that β is species specific (or sex specific for dimorphic species) is likely to be an adequate approximation for representative sample sizes. This is supported by variation in scaling trends between M and L at inter- and intraspecific levels (e.g. Green, Figuerola & King 2001). Comparisons of CIs are valid when made among groups sharing the same β value for M - L relationships, regardless of variation in growth rate between individuals, populations or sex. Other comparisons of CIs (e.g. between sex that lack a common morphogenetic pattern) are nonsensical.

According to the dimensional balance between volume V (closely related to M) and length L (i.e. $M \equiv V \propto L^3$), under 'isometry' the scaling exponent that relates M and L is 3. This provides a rationale for the formulation of the Fulton's index ($K = M/L^3$). In practice, β for M against L usually deviates from the predicted value of 3, owing to 'allometry'. Body condition should be estimated at no higher than the species level, which comprises the scaling effects owing to 'heterauxesis' and 'individual allomorphosis' (Gould 1966). 'Heterauxesis' (also called 'ontogenetic allometry' or 'growth allometry') refers to scaling during the growth of an individual, whilst 'allomorphosis' (or 'static allometry') is scaling among conspecifics at the same stage of determinate growth (usually adults), but varying in size. Although heterauxesis and allomorphosis require different types of data, both types of scaling are relevant to body condition as they are parts of a common morphogenetic phenomenon (Stern & Emlen 1999).

To enable a meaningful comparison between individuals of different sizes, a CI method must remove the effects of ontogenetic growth on the M - L relationship through standardisation. Unless some factor affects the well-being of animals at a specific growth stage (e.g. iguanas can become inefficient foragers when they reach a certain size, Wikelski, Carrillo & Trillmich 1997), mean condition scores should be equal for different age classes, and confirmation of this indicates that size and composition have been properly standardized (i.e. growth effects have been accounted for). Similarly, even if a species shows sexual size dimorphism, if the sexes have a similar body design (i.e. a similar morphogenesis), we should expect no differences in CIs (i.e. well-being) between sex.

Materials and methods

MATCHING THE SCALED MASS INDEX TO FIRST PRINCIPLES

The Scaled mass index (eqn 1) is based on a mathematical rearrangement of the simple power function $Y = \alpha X^\beta$ adapted for the study of body condition (see Appendix S1). The scaling exponent β is the most important parameter of interest (i.e. b_{SMA}) in eqn. 1, and should ideally be estimated from large data set (e.g. large reference populations covering the whole range of body size).

β is estimated by regression between M and L , following natural log transformation of the power equation (i.e. $\ln M = \ln \alpha + \beta \ln L$). Because of the natural M - L interdependence and their distinct scale of measurement, the scaled mass index relies on standardized major axis (SMA) regression to estimate the regression coefficient β ($\approx b_{SMA}$). SMA (also called reduced major axis), a form of model II regression, assumes variation in both Y and X variables (see below), and is recommended to ascertain the structural relationship between morphometric data, when the scaling exponent is the parameter of primary interest (Seim & Sæther 1983; LaBarbera 1989; Warton *et al.* 2006).

The Scaled mass index standardizes all individuals to the same L value and adjusts their body mass with that which they would have at their new L value in accordance to the scaling trend between M and L . In other words, it standardizes all individuals to the same growth phase. For the purpose of comparing condition between individuals, phenomena of body growth variation linked to morphogenesis (e.g. ontogeny or sexual size dimorphism) can be seen as shifts along the line (*sensu* Warton *et al.* 2006) described by the scaling exponent (estimated by b_{SMA}).

Thus, unlike conventional CI methods, the Scaled mass index is compatible with the following precepts underlying studies of condition (see Peig & Green 2009 and Table S1):

1. Growth is a multiplicative process of living material (Huxley 1932; Shea 1985), which makes nonlinear (power) models for M - L relationships more appropriate than linear ones.
2. The existence of allometry (i.e. departures from isometry), making empirical estimation of the scaling exponent necessary (Gould 1966; Schmidt-Nielsen 1984).
3. The specificity of the morphogenetic pattern determining the body proportions for a particular species, subspecies, or sex within a species (Conlon & Raff 1999; Hogan 1999), hence the need for a reliable estimate of the scaling exponent in each case.
4. The mutual interdependence of M and L driven by body growth, with both being inherent indicators of true size (i.e. body volume) (Thompson 1961).
5. The concomitant change of body composition with size owing to morphogenesis, such that only individuals at the same growth phase are strictly comparable (Kotiahlo 1999).

A CASE STUDY FRAMEWORK

We compared the performance of seven different CI methods using data from three species of small mammals collected in the Garraf Massif, northeast Spain from 2000 to 2003 for an ecotoxicological survey. Wood mice (*Apodemus sylvaticus* Linnaeus) ($n = 272$), Algerian mice (*Mus spretus* Lataste) ($n = 144$), and greater white-toothed shrews (*Crocidura russula* Hermann) ($n = 142$) were live-trapped in riparian habitats differing in pollution intensity. Two sites, 'Medium' and 'High' respectively, were affected by landfill leachates rich in heavy metals and organic pollutants from the largest dumping site in

Spain (Sánchez-Chardi *et al.* 2007, J. Peig unpublished data). A third riparian site unaffected by landfill leachates was selected as a 'reference' site.

For each individual, body mass (M) was measured with a Pesola® spring balance (± 0.1 g). Stomach contents, embryos, and parasites were also weighed using a digital balance (± 0.001 g) to correct final body mass. Four linear body measurements (L values) [body (from snout-to-anus), tail, left hind foot, and left ear length] were taken with a square calliper to the nearest millimetre. Each individual was sexed and the degree of sexual maturity determined by standard examination of reproductive structures (e.g. Díaz & Alonso 2003). Specimens were divided into three age classes (I, II or III, see Appendix S2 for details).

SELECTING MASS-LENGTH VARIABLES AND DATA SET

We examined bivariate linear correlations of M against L (both \ln -transformed), using original length variables or a combination of them following principal component analysis (see Appendix S3). As body length was the linear measurement most correlated with M it was selected as the L variable for all CI methods (Appendix S3). A Gaussian distribution of $\ln M$ and $\ln L$ was assumed based on the Kolmororov-Smirnov Z test ($P > 0.05$) and P-P plots.

Each of the seven methods was used to calculate a CI for all individuals. According to recommended procedures for computing \hat{M}_i and W_r , only the most reliable data were used to estimate key parameters (Murphy, Willis & Springer 1991; Peig & Green 2009). Hence, to calculate \hat{M}_i , we first performed bivariate ($\ln M$ - $\ln L$) plots to establish which set of individuals provided the most reliable estimate (b_{SMA}) of the scaling exponent β . For all species, the strength of fit improved when only individuals from the reference population were included: the strength of association ($r^2 \times 100$) between $\ln M$ and $\ln L$ increased from 84% (all individuals) to 87% (reference group) for wood mice, from 77% to 93% for Algerian mice, and from 35% to 47% for shrews. A likely reason is that animals inhabiting polluted areas are affected by toxins which disrupt growth trajectories, i.e. the true M - L relationship or β . The phenotypic variance in body length can be used as a measure of developmental instability in small mammals (Velickovic 2007). For each study species, the total variability in body length increased in parallel to pollution intensity, although differences were only significant in one species [variance in body length: $s^2_{\text{Ref.}} = 61.30$, $s^2_{\text{Med.}} = 84.12$, $s^2_{\text{High}} = 88.62$ ($n = 272$, $P = 0.277$) for wood mice; $s^2_{\text{Ref.}} = 37.88$, $s^2_{\text{Med.}} = 43.64$, $s^2_{\text{High}} = 44.96$ ($n = 144$, $P = 0.020$) for Algerian mice; and, $s^2_{\text{Ref.}} = 24.60$, $s^2_{\text{Med.}} = 35.65$, $s^2_{\text{High}} = 43.43$ ($n = 142$, $P = 0.157$) for shrews]. Thus, although \hat{M}_i was applied to the whole sample, the b_{SMA} was estimated using the reference sample. Likewise, parameters α and β required for W_r were obtained from the reference population.

RELATING CONDITION INDICES TO AGE, SEX AND LOCATION

Age class and sex were selected as natural factors causing differences in body size measured either as M and L . Using seven CI methods, we analyzed the differences in CIs between age classes and sex. CIs are also used to compare groups in different habitats or undergoing different treatments. Hence, we considered the influence of study site (i.e. pollution intensity) on CIs. A detailed interpretation of site/pollution effects on CIs will be given elsewhere.

For \hat{M}_i (Eq. 1), b_{SMA} was calculated from the reference group using RMA for Java v.1.19 software (Bohanak & van der Linde

2004). The arithmetic mean of body length from the whole population was taken as L_0 [98.22 mm ($n = 272$) for wood mice, 79.92 mm ($n = 144$) for Algerian mice, and 67.03 mm ($n = 142$) for shrews]. The Body mass index ($\text{BMI} = M/L^2$) and Fulton's index ($= M/L^3$) were computed using raw mass-length data. A multiplier of 10^3 for the BMI and 10^6 for K were included to give conventional units of kg m^{-2} and kg m^{-3} respectively. For Relative condition ($K_n = M/L^b$), a and b were estimated from the whole sample by OLS regression of the linearized power equation $\ln M = \ln a' + b \ln L$ [where a in K_n equals $e^{(\ln a')}$]. The same procedure was used to compute Relative mass (W_r), but in this case a and b were obtained from the reference population. The Residual index (R_i) was calculated from an OLS regression of $\ln M$ against $\ln L$. Finally, for ANCOVA, $\ln M$ was the dependent variable and $\ln L$ the covariate for each analyzed factor (García-Berthou 2001).

The resulting CI scores were analyzed with respect to age, sex and location with parametric (ANOVA and student's t -test) and non-parametric tests (Kruskal-Wallis H test and Mann-Whitney U test) to identify significant (at $\alpha < 0.05$) or marginal (at $0.05 < \alpha < 0.10$) effects. For parametric tests, the assumption of homoscedasticity was explored via the Levene test at $\alpha = 0.05$. For ANCOVA, we tested the critical assumption of parallelism in regression lines between factor levels at $\alpha = 0.05$ (García-Berthou 2001).

COMPLEMENTARY BIOSTATISTICAL ASSAYS

The different CI methods use different approaches for size standardisation and different assumptions about the relationships between M and L (see Table S1), which we tested. We assessed the underlying nature (linear vs. nonlinear) of the relationship between the selected M and L variables. We also examined the scaling exponents assumed by each CI method. Log-transformation usually normalizes data before using parametric models (linear regression, ANCOVA, etc.), and is also used to fit the conventional M - L relationship ($Y = \alpha X^\beta$) (Marshall, Jakob & Uetz 1999; García-Berthou 2001; Green 2001). Thus, the values of β can be compared using the regression coefficients b obtained from different linear models which take the general form $Y = a + bX$. For \hat{M}_i , the model used by SMA regression can be described as $Y_1 = a + b_{\text{SMA}} Y_2$ (McArdle 2003), where there is no independent variable and both variables (Y_1 - Y_2) are subjected to natural variability and measurement error. The slope b_{SMA} minimizes the areas of triangles (in the y - and x -axis direction from a $\ln M$ - $\ln L$ plot) between data points and the best-fitted line (McArdle 1988). R_i , K_n and W_r indices are based on the OLS (or Model I) regression $Y_i = a + b_{\text{OLS}} X_i + \varepsilon_i$, where there is an independent (X) variable with no natural variability or measurement error. It thus assumes an extremely asymmetrical relationship between Y and X , the slope b_{OLS} being drawn to minimize the sum of squares between data points and the line in the y -axis direction only. The ANCOVA model can be formulated as $Y_{ij} = a_i + b_{\text{pooled}}(X_{ij} + \bar{X}_i) + \varepsilon_{ij}$, combining features of model I regression (based also on the principle of least squares in the y plane) with one-way ANOVA (García-Berthou 2001). ANCOVA thus computes an average slope (b_{pooled}) of the regression coefficients defined by each factor level (a_i), whose lines are also fitted by assuming an independent covariate with no natural variability or measurement error. Finally, the BMI and K index use fixed values of β (referred to below as b_{presumed}).

We compared values of b_{SMA} , b_{OLS} , b_{pooled} , and b_{presumed} and considered their relationship with the above first principles. Furthermore, we evaluate the close relationship between some of the CI methods. Some methods produce CIs which are highly intercorrelated, yet there

remain important differences in their interpretation (see also Peig & Green 2009). Except for calculation of b_{SMA} , statistical analyses were performed with SPSS v.15 (SPSS Inc., Chicago, USA).

Results

COMPARING CONDITION INDICES IN RELATION TO AGE, SEX AND LOCATION

Relative age

Body mass and body length differed significantly between age classes for each species (Kruskal–Wallis H test: $P < 0.001$), except for marginal differences observed for body length in shrews ($P < 0.1$). As expected, both aspects of body size tended to increase with ontogeny for all species (Figure S1).

There were major differences in the relationship between CI and age class between methods. The Scaled mass index did not show age differences in CI for any species, whereas conventional methods usually indicated major differences between age classes (Table 1, Fig. 1). For all species, R_i , ANCOVA, BMI and K_n showed marked differences in CIs among age classes. Fulton's index (K) showed significant differences in wood mice, marginal differences in Algerian mice, but no differences in shrews. In contrast, Relative mass (W_r) showed major differences for shrews and wood mice, and marginal differences for Algerian mice.

Condition index values from conventional methods increased consistently with age class, with the exception of a

sequential decrease observed for K in wood mice (Fig. 1). The BMI and those CIs based on the principle of least squares along the y -axis (i.e. R_i , ANCOVA, K_n and W_r) showed the same trend with age class as did M and L (Figs 1 and S5). This indicates a systematic bias towards higher CI values in larger individuals.

Sex

Sexual size dimorphism was marked in wood mice (M : $\delta > \text{♀}$, $P < 0.01$; L : $\delta > \text{♀}$, $P < 0.01$) and Algerian mice (M : $\delta > \text{♀}$, $P < 0.05$; L : $\delta > \text{♀}$, $P > 0.10$), but not in shrews ($\delta \approx \text{♀}$, $P > 0.10$ for both M and L) (Figure S1). For all species, comparisons of CIs between sex were legitimate because the scaling relationship between M and L was the same for both sex, as shown by a lack of difference in slope ($P > 0.10$) when fitting a common log-log SMA regression. Thus, despite sexual differences in size, we assumed the same morphogenetic pattern for males and females.

The Scaled mass index showed no sexual differences for any species (Table 1, Fig. 2). In contrast, all conventional methods (except for the Fulton's index ' K ') showed significant or marginal effects of sex on CIs, for at least one species. These statistical differences were exclusively recorded in the size-dimorphic species (i.e. wood mice and Algerian mice, Figure S1). Even when differences between these two species were not significant, most conventional methods tended to inflate CIs for males, owing to their larger body size (Fig. 2).

Table 1. Effects (P -values) of age class, sex and location (population) on body condition estimates for three species of small mammals according to seven different methods. Bold type indicates a significant factor effect at $\alpha = 0.10$. See Table 2 and introduction for details of each method

Factors by species		\hat{M}_i	R_i	ANCOVA	BMI	K	K_n	W_r
<i>Relative age</i>								
Wood mouse	PA	0.741	< 0.001	< 0.001†*	< 0.001	0.022*	< 0.001	0.020
	NP	0.775	< 0.001	—	< 0.001	0.030	< 0.001	0.007
Algerian mouse	PA	0.368	< 0.001	< 0.001†*	< 0.001*	0.073	< 0.001*	0.093
	NP	0.471	< 0.001	—	< 0.001	0.104	< 0.001	0.126
Greater w-t shrew	PA	0.441	< 0.001	< 0.001	0.083	0.658	< 0.001	0.003
	NP	0.318	< 0.001	—	0.054	0.497	< 0.001	0.002
<i>Sex</i>								
Wood mouse	PA	0.792	0.099	0.086†	0.037	0.290	0.135	0.350
	NP	0.508	0.014	—	0.003	0.759	0.014	0.054
Algerian mouse	PA	0.450	0.119	0.119	0.096	0.345	0.105	0.356
	NP	0.982	0.263	—	0.244	0.787	0.261	0.803
Greater w-t shrew	PA	0.939	0.949	0.949	0.993	0.917	0.911	0.942
	NP	0.761	0.682	—	0.640	0.783	0.682	0.606
<i>Population</i>								
Wood mouse	PA	0.017*	0.013	0.014*	0.013*	0.025*	0.012*	0.013*
	NP	0.019	0.010	—	0.009	0.030	0.010	0.011
Algerian mouse	PA	0.045*	0.312*	< 0.001†*	0.219*	0.059*	0.202*	0.057*
	NP	0.109	0.570	—	0.591	0.194	0.571	0.189
Greater w-t shrew	PA	0.001	0.004	0.004	0.001	0.002	0.003	0.001
	NP	0.002	0.011	—	0.002	0.002	0.011	0.004

* $P < 0.05$ for the Levene's homoscedasticity test.

†Heterogeneity of slopes in the ANCOVA method at $\alpha = 0.05$.

PA, parametric tests (one-way ANOVA or Student t -test).

NP, non-parametric tests (Kruskal–Wallis H test or Mann–Whitney U test).

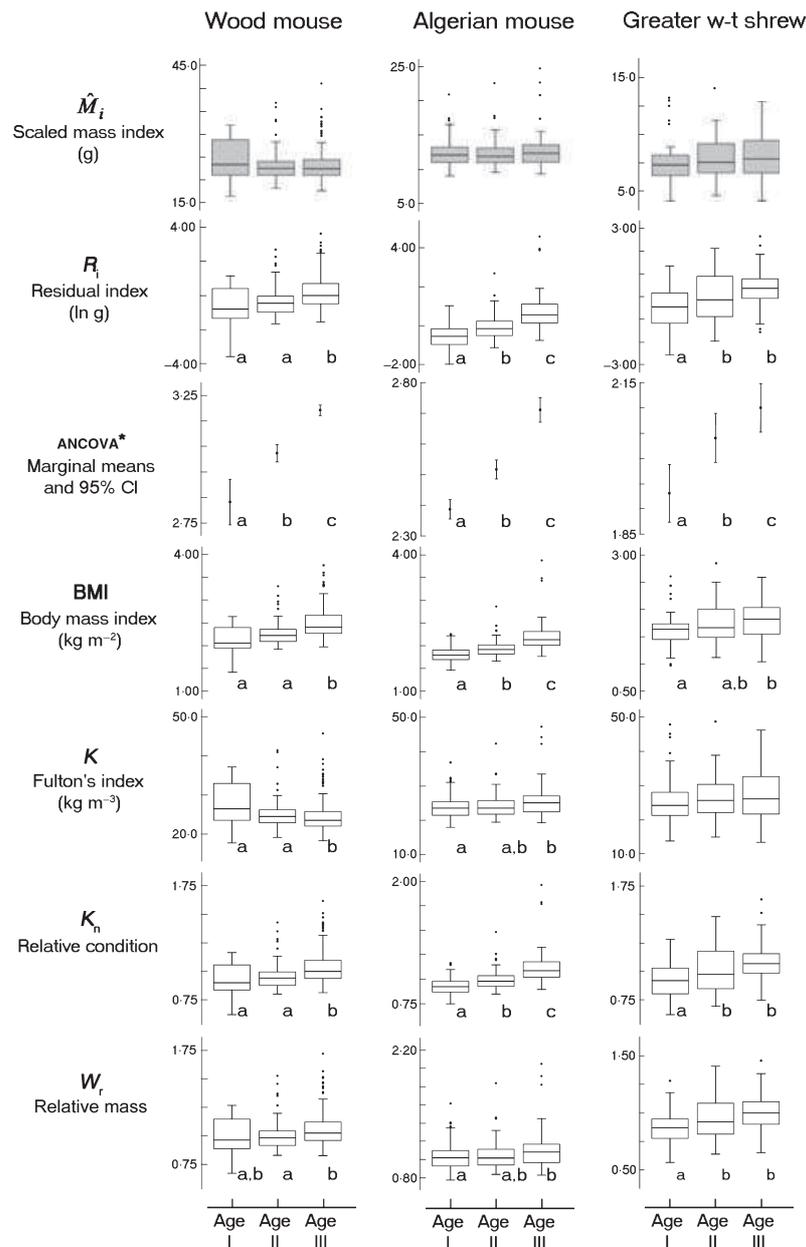


Fig. 1. Ontogenetic trends in body condition index (CI) estimated by seven different methods. There were no significant differences between age classes for \hat{M}_i . Conventional methods produced CIs that differed between age classes, with the exception of K for shrews (Table 1). Where appropriate, alphabetic characters (a,b,c) below each plot summarize pairwise post-hoc comparisons among age classes after Bonferroni correction (age classes not sharing letters were statistically different). CI values exceeding 1.5 times the interquartile range of boxplots are marked with black dots. *Marginal means and 95% CI for ANCOVA were evaluated at $\ln(\text{body length}) = 4.58$ for wood mice, 4.38 for Algerian mice, and 4.20 for shrews.

Location

According to \hat{M}_i , sampling location had significant effects on condition for all three species (Table 1). Conventional CIs for wood mice and shrews consistently varied with location. For Algerian mice, R_i , BMI and K_n failed to detect a difference between locations. In contrast, ANCOVA, K and W_r suggested there were differences in condition among populations (despite violations of statistical assumptions), in accordance with \hat{M}_i . Because Algerian mice were subjected to the same ecological disturbance as the other

two species, it seems likely that there was a real difference in condition between locations, which R_i , BMI and K_n failed to detect.

COMPARING ASSUMPTIONS ABOUT SCALING

The relationship between M and L was clearly nonlinear for all species (Appendix S4). The various CI methods relied on different values for the scaling exponent β (Table 2). For all three mammalian species, the various regression coefficients (b_{SMA} , b_{OLS} , b_{pooled}) from the linear models suggested an allo-

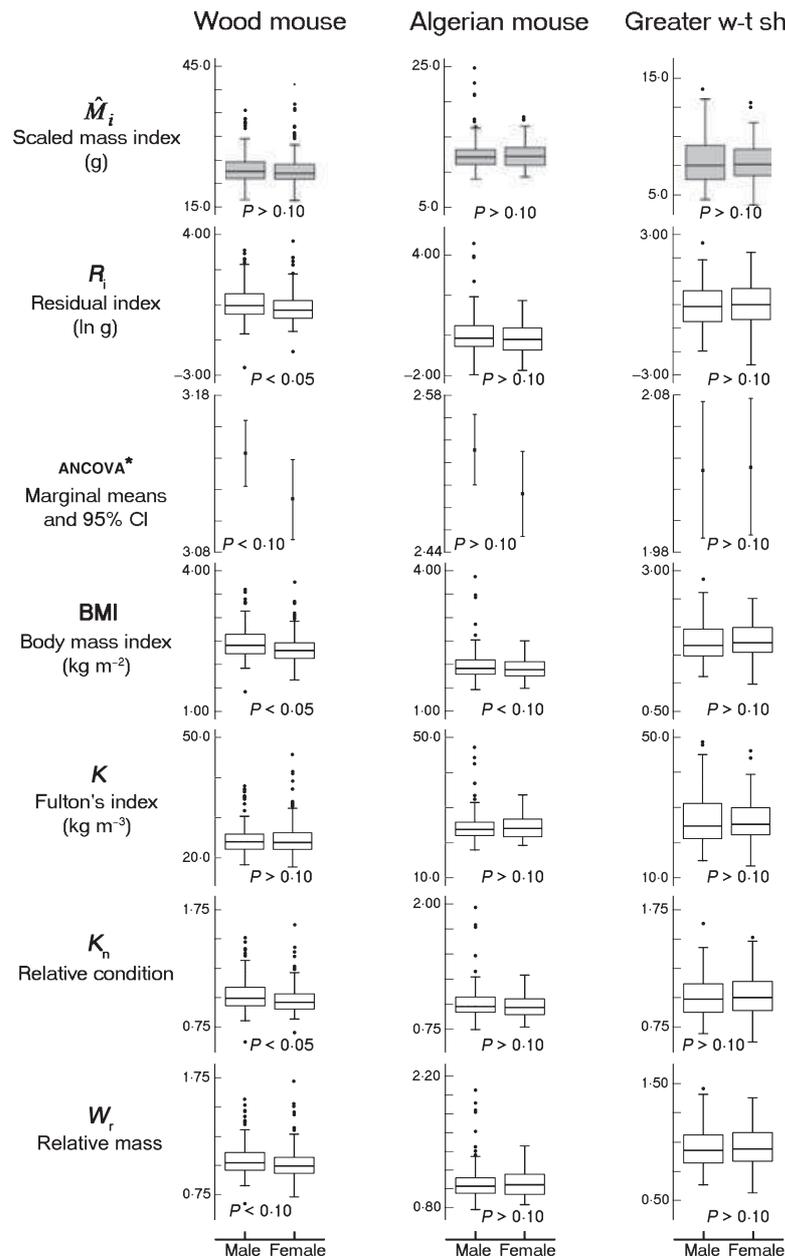


Fig. 2. Sexual trends in body condition index (CI) estimated by seven different methods. There were no significant differences between sex for \hat{M}_i even for size-dimorphic species (i.e. wood mice and Algerian mice). In contrast, most conventional methods yielded statistical differences in CIs between sex (at $\alpha = 0.10$), and sex-differences in CIs were male-biased in accordance with the pattern of size-dimorphism. CI values exceeding 1.5 times the interquartile range of boxplots are marked with black dots. *Marginal means and 95% CI for ANCOVA were evaluated at $\ln(\text{body length}) = 4.58$ for wood mice, 4.38 for Algerian mice, and 4.20 for shrews.

metric/power relationship between M and L . Although b values of 3 are expected if growth follows simple geometric rules, in reality β may deviate somewhat from this value because growth is not isometric. For instance, $b_{\text{SMA}} = 2.71$ appears to be most reliable approximation for β in wood mice (Table 2) because \hat{M}_i is the only CI independent of body size (i.e. independent of ontogeny and sex). The Fulton's index ' K ' assumes an exaggerated value of 3, thus underestimating the relative mass of larger animals and causing a progressive reduction in CIs with ageing and growth (Fig. 1). On the other hand, ' b ' values below 2.71 ($\cong \beta$) in wood mice inflate CIs in larger animals, producing higher values in

older individuals and in males (e.g. R_i , ANCOVA, BMI, K_n , W_r in Figs 1 and 2).

Likewise, in Algerian mice $b_{\text{SMA}} = 3.26$ is the most reliable β estimate, and lower values of b (< 2.1) led to consistently biased CIs in larger animals (i.e. R_i , ANCOVA, BMI and K_n , Figs 1 and 2). The level of bias was much weaker in methods using higher b values ($b_{\text{presumed}} = 3$ for K and $b_{\text{OLS}} = 3.03$ for W_r) and only weak age and sex effects were recorded for these methods (Table 1). Finally, in shrews, low ' b ' values (< 1.26) used by R_i , ANCOVA, BMI, K_n and W_r explained their systematic bias towards larger animals (Fig. 1). Only \hat{M}_i ($b_{\text{SMA}} = 2.68$) and K ($b_{\text{presumed}} = 3$) produced CIs

Table 2. Details of scaling exponents β used by different condition index (CI) methods in three species of small mammals. From the conventional mass-length relationship (i.e. $M = \alpha L^\beta$), the slope β can be estimated by linear models fitted to an $\ln M$ - $\ln L$ biplot: b_{SMA} is the regression coefficient of a standardized major axis (SMA) regression ($Y_1 = \alpha' + \beta Y_2$), b_{OLS} is the regression coefficient of an ordinary least squares (OLS) regression ($Y = \alpha' + \beta X$), and b_{pooled} is the regression coefficient of the ANCOVA model ($Y_{ij} = \alpha_i + \beta(X_{ij} + \bar{X}_i) + \varepsilon_{ij}$). b_{presumed} are fixed values of β and are not estimated from linear models. n is lower for methods in which a reference population was used to calculate β

CI method	Type of scaling exponent	Wood mouse			Algerian mouse			Greater w-t shrew		
		n	r^2	Value [95% CI]	n	r^2	Value [95% CI]	n	r^2	Value [95% CI]
Scaled mass index (\hat{M}_i)	b_{SMA}	97	0.75*	2.71 [2.43, 2.99]	30	0.88*	3.26 [2.83, 3.69]	47	0.22*	2.68 [1.97, 3.39]
Residual index (R_i)	b_{OLS}	272	0.70*	2.18 [2.01, 2.34]	144	0.59*	2.08 [1.79, 2.36]	142	0.12*	0.79 [0.44, 1.13]
ANCOVA method for										
Relative age	b_{pooled}	272	†	1.36 [1.11, 1.60]	144	†	1.13 [0.86, 1.39]	142	0.22*	0.63 [0.30, 0.97]
Sex	b_{pooled}	272	†	2.13 [1.96, 2.31]	144	0.60*	2.05 [1.77, 2.34]	142	0.11*	0.79 [0.44, 1.14]
Location	b_{pooled}	272	0.71*	2.18 [2.01, 2.35]	144	†	2.09 [1.80, 2.38]	142	0.18*	0.87 [0.53, 1.21]
Body mass index (BMI)	b_{presumed}	—	—	2.00	—	—	2.00	—	—	2.00
Fulton's index (K)	b_{presumed}	—	—	3.00	—	—	3.00	—	—	3.00
Relative condition (K_n)	b_{OLS}	272	0.70*	2.18 [2.01, 2.34]	144	0.59*	2.08 [1.79, 2.36]	142	0.12*	0.79 [0.44, 1.13]
Relative mass (W_r)	b_{OLS}	97	0.75*	2.34 [2.06, 2.61]	30	0.88*	3.03 [2.58, 3.58]	47	0.22*	1.25 [0.54, 1.96]

* $P < 0.05$ in the ANOVA test for fit of the linear model.

†Heterogeneity of slopes in the ANCOVA method at $\alpha = 0.05$.

CI, condition index.

independent of size variation. For this species, both methods performed equally well, suggesting that the true β may lie between $b_{\text{presumed}} = 3$ and $b_{\text{SMA}} = 2.68$. The M - L data for shrews may have been insufficient to evaluate a reliable β value, or morphogenetic pattern, as suggested by the particularly weak correlation between M and L (Appendix S3).

The linear regression models employed for \hat{M}_i , R_i , K_n and W_r were all significant (Table 2). However, five of nine ANCOVAs had violations of the critical parallelism assumption and/or homogeneity of variances (Tables 1 and 2). Furthermore, for ANCOVA, the slope (b_{pooled}) had to be recomputed for each explanatory factor analyzed, leading to repeated changes in the scaling exponent β from the same data set and reducing its biological meaning.

SIMILARITIES AND DIFFERENCES BETWEEN SOME CI METHODS

In some cases, the CIs produced by different methods were perfectly correlated (Appendix S5). However, these methods were not mathematically equivalent, as the CI values and the manner they should be interpreted differ among methods (Appendix S5).

Discussion

Decisions on the methods used for estimating body condition should be made cautiously and in compliance with theoretical precepts. As the aim of CIs based on M and L is to remove the growth effects on M with the ultimate goal of assessing an individual's health, CIs should be independent of age, sex and other natural causes of body size variation linked to morphogenesis. The pervasive effects of these intrinsic factors on CIs has been recognized as an important limitation of the Body mass index (BMI) in epidemiology (Gallagher *et al.* 1996; Prentice & Jebb 2001), but are usually ignored in ecological

studies. A major finding of our reappraisal is that the Scaled mass index \hat{M}_i accounted for ontogenetic variation or sexual dimorphism in body size, while all conventional methods failed to consistently standardize these key parameters. As we anticipated, all methods based on the principle of least squares in the y -plane (R_i , ANCOVA, K_n and W_r) were systematically biased towards larger individuals, despite the current acceptance of these methods in ecology or fisheries biology. In fact, in recent years the statistical mechanics have been the main concern of authors proposing particular CI methods (e.g. García-Berthou 2001; Freckleton 2002; Schulte-Hostedde *et al.* 2005), overlooking the importance of underlying natural factors on condition estimates as definitive proof of their suitability.

Although Fulton's index (K) has been subjected to much criticism (see Stevenson & Woods 2006), it performed better than more sophisticated techniques (e.g. R_i and ANCOVA), despite producing CIs that decreased with size in the wood mouse. K probably performed well because the scaling relationship it assumes ($M-L^3$) is closer to reality than those suggested by the biased slopes from OLS regression of M on L . The BMI also showed systematic bias in larger animals. Our results explain why the conventional methods often compel ecologists and epidemiologists to compare condition among individuals that are relatively homogeneous in body size, i.e. by separating ages or sex, thus circumventing undesired effects of growth on CIs (e.g. Deurenberg, Weststrate & Seidell 1991; Bister *et al.* 2000; Schulte-Hostedde, Millar & Hickling 2001; Velando & Alonso-Alvarez 2003). Otherwise, the basic premise of size-independence is likely to be broken (Gallagher *et al.* 1996; Green 2001). For broad ranges of body size, BMI and methods based on least squares in the y -axis automatically cause CIs to change with ontogeny or sexual size dimorphism (see also Blackwell 2002; Forsyth *et al.* 2005). These biases are likely to be reduced by using \hat{M}_i based on SMA regression. However, owing to its reliance

on log transformation to calculate b_{SMA} , even \hat{M}_i may be somewhat biased (Packard 2009).

Some reasons why conventional methods may produce misleading CIs were previously pointed out by Green (2001) and Peig & Green (2009). However, this current article represents a comprehensive reappraisal of available methods, based on empirical evidence and biostatistical precepts. The first precept is the existence of scaling and the nonlinear structural relationship between M and L . CI methods based on OLS linear models like R_i or ANCOVA can not account for the true M - L relationship, with increasing bias for unrepresentative sample sizes (e.g. covering small ranges of body size which lead to particularly low r^2 values between M and L). A second precept is the existence of allometry (i.e. departures from isometry), which requires the empirical estimation of the scaling exponent β as the best description of the true M - L structural relationship. Presupposing values of β (e.g. $b = 2$ in the BMI or $b = 3$ in the K index) cause a bias in CIs owing to the deviation from the true value of β . The third precept is the mutual interdependence between M and L measurements, and its implications for β . For these two interdependent variables, the true slope will lie somewhere between the extreme slopes fitted by OLS regression of $\ln M$ on $\ln L$ and the reverse regression of $\ln L$ on $\ln M$ (McArdle 2003), the exact position depending on the relative importance of each variable as a descriptor of true body size.

The SMA regression is an *error in variables* model which best accounts for the natural M - L (or L - M) interdependence while accounting for the different scale of measurement (Warton *et al.* 2006; Peig & Green 2009). For every species, b_{SMA} values not only solved the confounding effects of growth, but they were also more in accordance with scaling theory and previous literature (reviewed by Green 2001), which suggest that the scaling exponents relating M and body length for mammals lie between 2.5 and 3.2 (Table 2). As the information on body size is partitioned between mass and body length, this range [2.5, 3.2] could serve as a guideline to identify reliable estimates of β in mammals. As there is no universal solution for finding the line of best fit between two interdependent variables (McArdle 1988), the use of reference or standard populations can help to define a suitable estimate of the β . For this reason, we used only data from uncontaminated sites to estimate b_{SMA} for the Scaled mass index. However, even when b_{SMA} was computed for the entire study sample, our \hat{M}_i estimates still performed better than the R_i , ANCOVA, BMI, K_n and W_r (results not shown). The only noteworthy change was that the differences in \hat{M}_i between age classes became significant for Algerian mice ($P < 0.05$).

The use of OLS regression to fit a line to M - L data is often defended because the residual scores are not correlated with body size or, in particular, with the length measurement L (e.g. Jakob, Marshall & Uetz 1996; Marshall, Jakob & Uetz 1999; Schulte-Hostedde, Millar & Hickling 2001). Nonetheless, this is an invalid justification, based on the incorrect assumption that L is the only indicator of body size and is an error-free measure. R_i and other methods based on least squares in the y -plane will tend to underestimate the regres-

sion coefficient (Table 2), being more attenuated (closer to zero) as absolute values of the correlation coefficient decrease (Seim & Sæther 1983; White & Seymour 2005; Arnold & Green 2007). Regression coefficients from conventional methods were well below the isometric value of 3 (with the exception of W_r in the Algerian mouse), with values below 0.8 in shrews. β estimates below the true M - L slope lead CIs to increase with increasing true body size, while β estimates above the true slope make CIs decrease. Very low values of regression coefficients should be taken as an early warning that the model is inappropriate for scaling and CI studies (see also LaBarbera 1989). On the other hand, SMA slopes can be more biased than OLS slopes when error rates in M are more than three times those in L , and when the correlation between M and L is high (> 0.9), the choice of regression method makes little difference (McArdle 1988; Arnold & Green 2007). Furthermore, all major statistical packages are based on OLS techniques, making them more suitable for multivariate analyses.

Despite the popular use of BMI in humans (taking height as L), it has rarely been applied to animals (Stevenson & Woods 2006). The BMI assumes $\beta = 2$ instead of 3 as predicted by isometry, because $b_{\text{OLS}} = 1.94$ for a regression of mass on height in humans (Khosla & Lowe 1967). In fact, our results and scaling theory suggest that β approximates much more to a value of 3 than to 2, indicating that the BMI will be dependent on size, as illustrated by the way it varied with ontogeny and sexual size dimorphism. We expect that \hat{M}_i is more appropriate for human studies than BMI.

In conclusion, we have presented solid evidence that conventional methods neither satisfy biological precepts concerning M - L relationships nor control successfully for variation in body size, often making them unreliable predictors of an individual's body condition. Furthermore, from a pragmatic viewpoint, some methods are very sensitive to statistical assumptions, especially ANCOVA which failed to provide reliable results in most of our analyses. Our results, together with those of Peig & Green (2009), suggest that currently fashionable methods (e.g. R_i and ANCOVA) can produce unreliable results.

In contrast, the Scaled mass index emerges as a powerful tool for condition studies from mass-length data. This method provides empirical results in accordance with the theoretical precepts identified for studies of condition, and has previously been validated with data on body components such as fat and protein (Peig & Green 2009). Being simple, reliable, comprehensive and easily interpretable, we encourage ecologists to implement this method in future studies. It is ideal both for comparing individuals of a single population, and for comparing CIs between populations or even between studies on the same species.

Acknowledgements

We thank Josep Lluís Parra, José Domingo Rodríguez-Tejedor and Jacint Nadal for valuable comments on earlier drafts. We also appreciate the technical assistance from the Gavà Museum. This work was funded by the Spanish

(project #: BOS 2000-0567) and Catalan Government (project #: 2001SGR-00090 and 2001ACOM-00009).

References

- Albrecht, G.H., Gelvin, B.R. & Hartman, S.E. (1993) Ratios as a size adjustment in morphometrics. *American Journal of Physical Anthropology*, **91**, 441–468.
- Ardia, D.R. (2005) Super size me: an experimental test for the factors affecting lipid content and the ability of residual body mass to predict lipid stores in nestling European starlings. *Functional Ecology*, **19**, 414–420.
- Arnold, T.W. & Green, A.J. (2007) On the allometric relationship between size and composition of avian eggs: a reassessment. *Condor*, **109**, 705–714.
- Bister, T.J., Willis, D.W., Brown, M.L., Jordan, S.M., Neumann, R.M., Quist, M.C. & Guy, C.S. (2000) Proposed standard weight (Ws) and standard length categories for 18 warmwater nongame and riverine fish species. *North American Journal of Fisheries Management*, **20**, 570–574.
- Blackwell, G.L. (2002) A potential multivariate index of condition for small mammals. *New Zealand Journal of Zoology*, **29**, 195–203.
- Bohanak, A. & van der Linde, K. (2004) RMA: software for reduced major axis regression, Java version (<http://www.bio.sdsu.edu/pub/andy/RMA.html>).
- Calder, W.A. (1984) *Size, function and life history*. Harvard University Press, Cambridge.
- Conlon, I. & Raff, M. (1999) Size control in animal development. *Cell*, **96**, 235–244.
- Deurenberg, P., Weststrate, J.A. & Seidell, J.C. (1991) Body mass index as a measure of body fatness: age- and sex-specific prediction formulas. *British Journal of Nutrition*, **65**, 105–114.
- Diaz, M. & Alonso, C.L. (2003) Wood mouse *Apodemus sylvaticus* winter food supply: density, condition, breeding and parasites. *Ecology*, **84**, 2680–2691.
- Forsyth, D.M., Duncan, R.P., Tustin, K.G. & Gaillard, J.M. (2005) A substantial energetic cost to male reproduction in a sexually dimorphic ungulate. *Ecology*, **86**, 2154–2163.
- Freckleton, R.P. (2002) On the misuse of residuals in ecology: regression of residuals vs. multiple regression. *Journal of Animal Ecology*, **71**, 542–545.
- Gallagher, D., Visser, M., Sepúlveda, D., Pierson, R.N., Harris, T. & Heymisfield, S.B. (1996) How useful is body mass index for comparison of body fatness across age, sex, and ethnic groups? *American Journal of Epidemiology*, **143**, 228–239.
- García-Berthou, E. (2001) On the misuse of residuals in ecology: testing regression residuals vs the analysis of covariance. *Journal of Animal Ecology*, **70**, 708–711.
- Gould, S.J. (1966) Allometry and size in ontogeny and phylogeny. *Biological Reviews*, **41**, 587–640.
- Green, A.J. (2001) Mass/length residuals: measures of body condition or generators of spurious results? *Ecology*, **85**, 1473–1483.
- Green, A.J., Figuerola, J. & King, R. (2001) Comparing interspecific and intraspecific allometry in the Anatidae. *Journal of Ornithology*, **142**, 321–334.
- Hogan, B.L.M. (1999) Morphogenesis. *Cell*, **96**, 225–233.
- Hoppeler, H. & Weibel, E.R. (2005) Scaling functions to body size: theories and facts. *Journal of Experimental Biology*, **208**, 1573–1574.
- Huxley, J.S. (1932) *Problems of relative growth*. Dover Publications Inc., New York.
- Jakob, E.M., Marshall, S.D. & Uetz, G.W. (1996) Estimating fitness: a comparison of body condition indices. *Oikos*, **77**, 61–67.
- Khosla, T. & Lowe, C.R. (1967) Indices of obesity derived from body weight and height. *British Journal of Preventive and Social Medicine*, **21**, 122–128.
- Kotiaho, J.S. (1999) Estimating fitness: comparison of body condition indices revisited. *Oikos*, **87**, 399–400.
- LaBarbera, M. (1989) Analyzing body size as a factor in ecology and evolution. *Annual Review of Ecology and Systematics*, **20**, 97–117.
- LeCren, E.D. (1951) The length–weight relationship and seasonal cycle in gonad weight and condition on the perch (*Perca fluviatilis*). *Journal of Animal Ecology*, **20**, 201–219.
- Marshall, S.D., Jakob, E.M. & Uetz, G.W. (1999) Re-estimating fitness: can scaling issues confound condition indices? *Oikos*, **87**, 401–402.
- McArdle, B.H. (1988) The structural relationship: regression in biology. *Canadian Journal of Zoology*, **66**, 2329–2339.
- McArdle, B.H. (2003) Lines, models, and errors: regression in the field. *Limnology and Oceanography*, **48**, 1363–1366.
- Murphy, B.R., Willis, D.W. & Springer, T.A. (1991) The relative weight index in fisheries management: status and needs. *Fisheries*, **16**, 30–38.
- Packard, G.C. (2009) On the use of logarithmic transformations in allometric analyses. *Journal of Theoretical Biology*, **257**, 515–518.
- Packard, G.C. & Boardman, T.J. (1999) The use of percentages and size-specific indices to normalize physiological data for variation in body size: wasted time, wasted effort? *Comparative Biochemistry and Physiology A*, **122**, 37–44.
- Peig, J. & Green, A.J. (2009) New perspectives for estimating body condition from mass/length data: the scaled mass index as an alternative method. *Oikos*, **118**, 1883–1891.
- Prentice, A.M. & Jebb, A. (2001) Beyond body mass index. *Obesity Reviews*, **2**, 141–147.
- Roth, V.L. & Mercer, J.M. (2000) Morphometrics in development and evolution. *American Zoologist*, **40**, 801–810.
- Sánchez-Chardi, A., Peñarroja-Matutano, C., Oliveira-Ribeiro, C.A. & Nadal, J. (2007) Bioaccumulation of metals and effects of a landfill in small mammals. Part II. The wood mouse, *Apodemus sylvaticus*. *Chemosphere*, **70**, 101–109.
- Schmidt-Nielsen, K. (1984) *Scaling: why is animal size so important*. Cambridge University Press, Cambridge.
- Schulte-Hostedde, A.I., Millar, J.S. & Hickling, G.J. (2001) Evaluating body condition in small mammals. *Canadian Journal of Zoology*, **79**, 1021–1029.
- Schulte-Hostedde, A.I., Zinner, B., Millar, J.S. & Hickling, G.J. (2005) Restitution of mass-size residuals: validating body condition indices. *Ecology*, **86**, 155–163.
- Seim, E. & Sæther, B.E. (1983) On rethinking allometry: which regression model to use? *Journal of Theoretical Biology*, **104**, 161–168.
- Serrano, E., Alpizar-Jara, R., Morellet, N. & Hewison, A.J.M. (2008) A half a century of measuring ungulate body condition using indices: is it time for a change. *European Journal of Wildlife Research*, **54**, 675–680.
- Shea, B.T. (1985) Bivariate and multivariate growth allometry: statistical and biological considerations. *Journal of Zoology*, **206**, 367–390.
- Stern, D.L. & Emlen, D.J. (1999) The developmental basis for allometry in insects. *Development*, **126**, 1091–1101.
- Stevenson, R.D. & Woods, W.A. (2006) Condition indices for conservation: new uses for evolving tools. *Integrative and Comparative Biology*, **46**, 1169–1190.
- Thompson, D.W. (1961) *On growth and form*. Cambridge University Press, Cambridge.
- Velando, A. & Alonso-Alvarez, C. (2003) Differential body condition regulation by males and females in response to experimental manipulations of brood size and parental effort in the blue-footed booby. *Journal of Animal Ecology*, **72**, 846–856.
- Velickovic, M. (2007) Measures of the developmental stability, body size, and body condition in the black-striped mouse (*Apodemus agrarius*) as indicators of a disturbed environment in northern Serbia. *Belgian Journal of Zoology*, **137**, 147–156.
- Warton, D.I., Wright, I.J., Falster, D.S. & Westoby, W. (2006) Bivariate line-fitting methods for allometry. *Biological Reviews*, **81**, 259–291.
- White, C.R. & Seymour, R.S. (2005) Allometric scaling of mammalian metabolism. *Journal of Experimental Biology*, **208**, 1611–1615.
- Wikelski, M., Carrillo, V. & Trillmich, F. (1997) Energy limits to body size in a grazing reptile, the Galapagos marine iguana. *Ecology*, **78**, 2204–2217.

Received 20 January 2010; accepted 14 June 2010

Handling Editor: Ryan Calsbeek

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Figure S1. Size variation in relation to age and sex.

Table S1. Desirable properties for an ideal CI.

Appendix S1. Mathematical development of the Scaled mass index.

Appendix S2. Determining age classes for small mammals.

Appendix S3. Selecting the most suitable linear body measurement for CIs.

Appendix S4. Nonlinearity of the relationship between body mass and length.

Appendix S5. Similarities and differences between some CI methods.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

FIGURE S1. Size variation in relation to age and sex.

Ontogenetic trends (A) and sex differences (B) in body mass ' M ' and body length ' L ' as intrinsic measures of size variation for three small mammal species. Size differences among age classes were all significant (Kruskal-Wallis test: $P < 0.001$), except for body length in shrews which was marginally significant ($P < 0.10$). Alphabetic characters (a,b,c) below each plot summarise pairwise post-hoc comparisons among age classes after Bonferroni correction (age classes not sharing letters were statistically different at $\alpha = 0.016$). Male-biased size dimorphism was marked in wood mice (Mann-Whitney test: $P < 0.01$) and Algerian mice ($P < 0.05$ for M), but not in shrews ($P > 0.10$ for both M and L). Values exceeding 1.5 times the interquartile range of boxplots are marked with black dots.

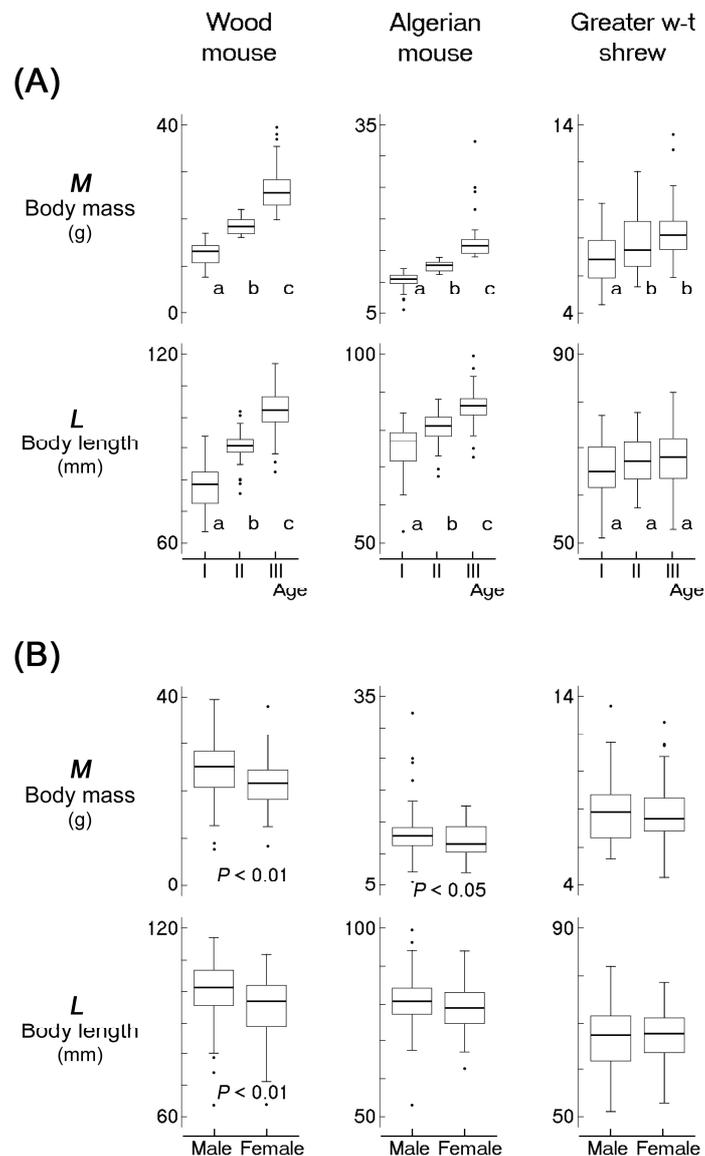


TABLE S1. Summary of desirable properties for an ideal condition index method based on mass M and length L , and the degrees to which different methods show these properties.

Desirable attribute	\hat{M}_i	R_i	ANCOVA	BMI	K	K_n	W_r
Assumes nonlinear (power) M - L relationship	Yes	No	No	Yes	Yes	Yes	Yes
Allows specific estimation of scaling component	Yes	Yes	Yes	No (assumed to be 2)	No (assumed to be 3)	Yes	Yes
Accepts that both M and L are size indicators	Yes	No	No	No	No	No	No
Assumes error exists in both M and L	Yes	No (Only in M)	No (Only in M)	No error structure	No error structure	No (Only in M)	No (Only in M)
Facilitates comparison in CI between studies	Yes	No	No	Yes	Yes	No	Yes
Produces CI in user-friendly units	Yes [g]	No [log _e g]	No units	No [kg m ⁻²]	No [kg m ⁻³]	No units	No units

Notes: CI methods are: Scaled Mass Index (\hat{M}_i), the Residual index (R_i , the residuals from an OLS regression of $\ln M$ against $\ln L$), ANCOVA, Quételet's index ($BMI = M/L^2$), Fulton's index ($K = M/L^3$), Relative condition (K_n , computed as the observed individual mass M_i divided by the predicted mass $M_i^* = a L_i^b$, where 'a' and 'b' are determined by OLS regression of $\ln M$ against $\ln L$), and Relative mass (W_r , where 'a' and 'b' above are determined from a reference population). More details of the characteristics of each method are found in the main text and in Peig and Green 2009.

Appendix S1. Mathematical development of the Scaled mass index in relation to theoretical precepts.

The Scaled mass index (Eq. 1 in the main text) is based principally on the formulation proposed by Lleonart, Salat and Torres (2000) for comparative morphology. However, while ‘size’ and ‘shape’ are the two properties of interest for most classic morphological studies (i.e. when relating variables of equal dimension [e.g. L_1 - L_2 , L_1^2 - L_2^2 , etc, where L denotes length]), ‘size’ and ‘density’ are ultimately those of most interest for developing CIs from body mass M and a linear body measurement L . In other words, as body growth involves variation in body size and body composition (Huxley 1932; Calder 1984), the construction of CIs from two external measurements like M and L implies, ideally, the standardisation of both body volume ‘ V ’ (i.e. true size in three dimensional space) and body density ‘ ρ ’ (i.e. body composition), according to the dimensional adjustment or scaling ($M \cong V \propto L^3$) and the principle of mass conservation ($\rho = M/V$) (Moya-Laraño *et al.* 2008). More specifically, since density is a physical property of a pure substance which is independent of its shape or volume, it is suitable to use the term ‘average body density’ to refer to the mass per unit volume of a biological entity made up of heterogeneous material.

Hence, following the steps of Lleonart *et al.* (2000), from the equation $M = aL^b$ relating body mass (M) to a linear body measurement (L) in an animal population or in an entire species, if ‘ b ’ contains the dimensional balance of changes in size and composition throughout growth (i.e. the morphogenetic pattern), then the information about body composition (e.g. fat, protein, etc,..) which is independent of growth must be gathered in the parameter ‘ a ’, and its value depends only on the characteristic body plan for that population or species. The body composition (a_i) of an individual i could be written as $a_i = a e^{\varepsilon_i}$ so that it is determined both by its belonging to a specific population or species (common factor in body composition or average density), denoted by a , and by its physiological or nutritional status (i.e. its particular body composition or average density), indicated as e^{ε_i} (this multiplicative error term is assumed to follow log-normality). Since particular features of body composition (e.g. nutritional status or the relative contribution of different components to final body mass) are transitory attributes which have distinct implications for animals of different size, one can assume

for convenience that the individual keeps constant the particular factor of composition e^{ε_i} (and its implications for individual health status) throughout its growth in order to compare such attribute among individuals of different size (see Lleonart et al. 2000 for more details).

With this assumption, if $M = a_i L^b$ expresses the ‘individual curve of morphogenetic growth’ for a particular animal with a pair of M - L measurements, the allometric equation $M = a L^b e^{\varepsilon_i}$ would express the ‘species curve of morphogenetic growth’ (by substituting a_i for $a e^{\varepsilon_i}$, where a describes the species-specific or common factor of body composition). The predicted body mass \hat{M}_i when the linear body measure attains a particular value ($L = L_0$) would be $\hat{M}_i = a L_0^b e^{\varepsilon_i}$ (Eq. 2). Note that L_0 denotes a particular point or stage in the morphogenetic pattern (i.e. along the curve) between M and L .

Similarly, from the species curve of morphogenesis $M = a L^b e^{\varepsilon_i}$, the hypothetical curve for an individual with M_i - L_i values could be represented as $e^{\varepsilon_i} = M_i / a L_i^b$ (Eq. 3). Substituting the eq. 3 into the eq. 2 we obtain $\hat{M}_i = a L_0^b (M_i / a L_i^b)$, that is, the predicted mass \hat{M}_i of the individual i for a given point (L_0) in the M - L relationship (i.e. for a given stage of morphogenesis). The algebraic rearrangement of this equation yields the Scaled mass index of condition \hat{M}_i (Eq. 1 in the main text), which is independent of a (common factor in body composition), and where b is the only estimate of interest to elucidate the scaling exponent β . Assuming biological meaning in β , it should ideally be estimated from the best possible dataset (e.g. ideally large reference populations covering the whole range of body size) as they may provide the best descriptor of the morphogenetic pattern guiding the body proportions.

The final formula (i.e. Eq. 1 in the main text) standardizes body mass at a fixed value of a linear body measurement based on the scaling relationship between mass and length. Consequently, for the purposes of comparing the Scaled mass index of condition between individuals differing in size, phenomena of body growth variation linked to morphogenesis (e.g. ontogeny and sexual size dimorphism) can be seen as shifts along the line (*sensu* Warton et al. 2006) described by the scaling exponent β .

References

- Calder, W.A. (1984) *Size, function and life history*. Harvard University Press. Cambridge.
- Huxley, J.S. (1932) *Problems of relative growth*. Dover Publications, Inc. New York.
- Leonart, J., Salat, J. & Torres, G.J. (2000) Removing allometric effects of body size in morphological analysis. *Journal of Theoretical Biology*, 205, 85–93.
- Moya-Laraño, J., Macías-Ordóñez, R., Blanckenhorn, W.U. & Fernández-Montraveta, C. (2008) Analysing body condition: mass, volume or density?. *Journal of Animal Ecology*, 77, 1099–1108.
- Warton, D.I., Wright, I.J., Falster, D.S. & Westoby, W. (2006) Bivariate line-fitting methods for allometry. *Biological Reviews*, 81, 259–291.

APPENDIX S2. Determining age classes for small mammals.

Specimens were divided into three age classes (I, II or III by combining morphometric data and traits of sexual activity for wood mice (e.g. Fernandez, Dunstone & Evans 1999; Erry *et al.* 2000; Lopes *et al.* 2002), via body size for Algerian mice (Nunes, da Luz Mathias & Crespo 2001), and according to the state of teeth abrasion for shrews (modified from Vesmanis & Vesmanis 1979). For wood mice, males of ≤ 17 g and females ≤ 16 g made up class I (juveniles), immature males between 17–22 g and immature females between 16–21 g were class II (subadults), and males ≥ 22 g and females ≥ 21 g or mature individuals weighing slightly less were class III (adults). For the Algerian mouse, we were unable to distinguish between subadults and adults from morphometry combined with reproductive traits, and only four juveniles (weighing < 9 g) were captured. Thus, we used arbitrary ranges of body sizes for each sex by examining bivariate plots of M against body length as a reasonable approximation of relative age, which allowed us to balance the sample among age classes. Class I corresponded to males of ≤ 12.5 g and females ≤ 11 g, class II to males of 12.5–14 g and females of 11–13 g, and class III to males ≥ 14 g and females ≥ 13 g. For the shrew, class I corresponded to state 1 and 2 of teeth abrasion (see Vesmanis & Vesmanis 1979), class II to state 3, and class III to states 4 and 5. Animals were handled in accordance with the established techniques and guidelines of the Animal Welfare Committee of the University of Barcelona.

References

- Erry, B.V., Macnair, M.R., Meharg, A.A. & Shore, R.F. (2000) Arsenic contamination in wood mice (*Apodemus sylvaticus*) and bank voles (*Clethrionomys glareolus*) on abandoned mine sites in southern Britain. *Environmental Pollution*, 110, 179–187.
- Fernandez, F.A.S., Dunstone, N. & Evans, P.R. (1999) Density-dependence in habitat utilisation by wood mice in a Sitka spruce successional mosaic: the roles of immigration, emigration, and variation among local demographic parameters. *Canadian Journal of Zoology*, 77, 397–405.

- Lopes, P.A., Viegas-Crespo, A.M., Nunes, A.C., Pinheiro, T., Marques, C., Santos, M.C. & Mathias, M.D.L. (2002) Influence of age, sex and sexual activity on trace element levels and antioxidant enzyme activities in field mice (*Apodemus sylvaticus* and *Mus spretus*). *Biological Trace Element Research*, 85, 227–239.
- Nunes, A.C., da Luz Mathias, M. & Crespo, A.M. (2001) Morphological and haematological parameters in the Algerian mouse (*Mus spretus*) inhabiting an area contaminated with heavy metals. *Environmental Pollution*, 113, 87–97.
- Vesmanis, I. & Vesmanis, A. (1979) Ein Vorschlag zur einheitlichen Altersabstufung bei Wimperspitzmäusen (Mammalia: Insectivora: Crocidura). *Bonner Zoologische Beiträge*, 30, 7–13.

Appendix S3. Selecting the most suitable linear body measurement for estimating body condition based on the relationship with body mass.

The first step to compute the Scaled mass index is to select the linear body measurement L best correlated with body mass M (Peig & Green 2009). Since many authors advocate use of L variables derived from principal component analysis (PCA) performed on several measures of length (e.g. Blackwell 2002; Schulte-Hostedde et al. 2005), we performed a PCA on body length, ear length, foot length and tail length (Table S3-A). We then conducted bivariate correlations between body mass M and L variables (individually or combined by PCA) (Table S3-B). As shown, a single linear measurement like body length was better correlated with mass than any other single L variable or their combination by PCA, indicating that the use of PCA should not be taken as a rule of thumb when estimating condition from mass-length data (see Peig & Green 2009 for more discussion).

TABLE S4-A. Principal component analysis conducted on four morphometric measurements (raw data) in three species of small mammals. The percent variance (eigenvalues) and factor loadings (eigenvectors) of the two first principal factors (F1, F2) are shown. All variables were normally distributed, except for tail length in the wood mouse.

	Wood mouse		Algerian mouse		Greater white-toothed shrew	
	F ₁	F ₂	F ₁	F ₂	F ₁	F ₂
Percent variance	47.0	23.3	57.8	18.8	43.4	24.8
Factor loadings						
Body length	0.83	-0.08	0.81	0.11	0.29	0.91
Tail length	0.43	0.86	0.78	-0.45	0.70	0.15
Foot length	0.76	-0.03	0.77	-0.25	0.80	-0.17
Ear length	0.66	-0.43	0.67	0.69	0.72	-0.33

TABLE S4-B. Linear correlations (r^2) between body mass and linear body measurements (individually or combined by PCA) in three species of small mammals. All variables were ln-transformed. * $P < 0.05$, ** $P < 0.01$].

	Body mass		
	Wood mouse	Algerian mouse	Greater w-t shrew
Body length	0.71**	0.60**	0.11**
Tail length	0.03*	0.17**	0.02
Foot length	0.19**	0.16**	0.01
Ear length	0.09**	0.17**	0.00
F1	0.56**	0.16**	0.03
F2	0.00	0.01	0.11**

References

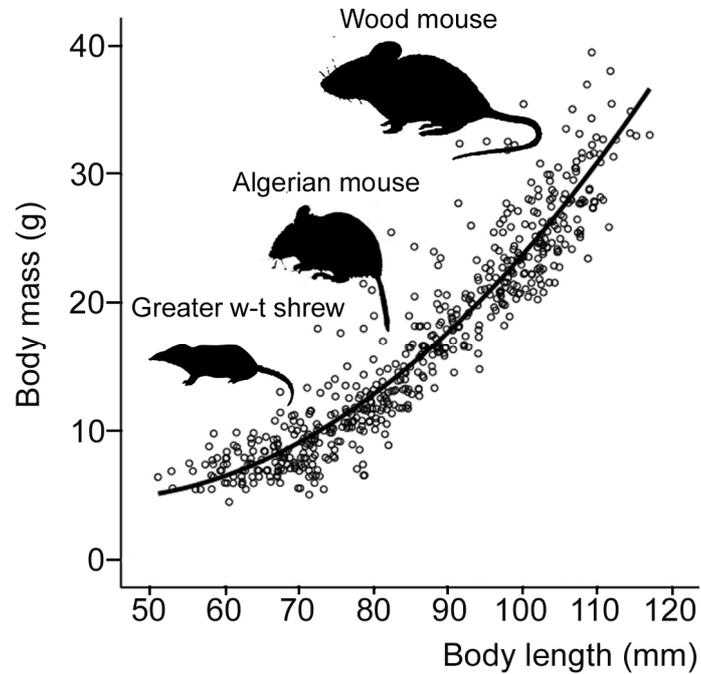
- Blackwell, G.L. (2002) A potential multivariate index of condition for small mammals. *New Zealand Journal of Zoology*, 29, 195–203.
- Peig, J. & Green, A.J. (2009) New perspectives for estimating body condition from mass/length data: the Scaled mass index as an alternative method. *Oikos*, 118, 1883–1891.
- Schulte-Hostedde, A.I., Zinner, B., Millar, J.S. & Hickling, G.J. (2005) Restitution of mass-size residuals: validating body condition indices. *Ecology*, 86, 155–163.

APPENDIX S4. Non-linearity of the relationship between body mass (M) and body length (L).

Usually, drawing a straight line through the bivariate (M - L) scatter provides a good visual and statistical fit (e.g., Schulte-Hostedde *et al.* 2005). Here, the linear function $M = \alpha + \beta L$ also provided a significant fit to the data ($r^2 = 0.66$ for the wood mouse, $r^2 = 0.51$ for the Algerian mouse, and $r^2 = 0.12$ for the shrew), but makes no biological sense given the existence of scaling. The non-linear relationship between M and L is illustrated by significant positive correlations between the ratio M/L and L for the mice species (Pearson's $r > 0.49$ and $P < 0.001$), and by the fact that the power function $M = \alpha L^\beta$ provides a better fit to the data ($r^2 = 0.70$ for the wood mouse, $r^2 = 0.59$ for the Algerian mouse, and $r^2 = 0.12$ for the shrew).

The nonlinearity between M and L is more obvious when the range of body sizes under study is increased. The three small mammal species differ in body size but share a similar 'mouse-type' general body design, and an interspecific plot showed a clear curvilinear trend between M and L (Fig. S6). This curvilinear relationship appears to correspond to the biomechanical and functional constraints posed by growth processes, with phylogenetic effects having minor importance for these species with a similar body design. After linearization of the interspecific M - L relationship by log transformation, the β estimate approached the value of 3 predicted from the simple expected dimensional adjustment for M - L variables ($b_{\text{SMA}} = 2.81$, 95% confidence interval [2.72, 2.89] or $b_{\text{OLS}} = 2.62$, 95% confidence interval [2.53, 2.70]). It is also noteworthy how this b_{SMA} value lies within the range of those recorded for each individual species, whereas this b_{OLS} value is much greater than those recorded for individual species (see Table 2 in the main text). Unlike SMA, OLS consistently underestimates β for the true scaling relationship, and the degree of underestimation is exacerbated as the range of body sizes is decreased.

FIG. S6. Interspecific plot showing the curvilinear relationship between body mass (M) and body length (L) for three mammal species with similar 'mouse-type' body design.



References

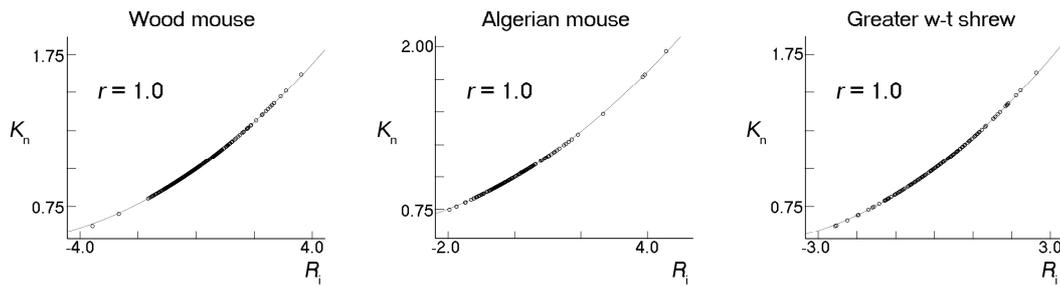
Schulte-Hostedde, A.I., Zinner, B., Millar, J.S. & Hickling, G.J. (2005) Restitution of mass-size residuals: validating body condition indices. *Ecology*, 86, 155–163.

Appendix S5. Similarities and differences between some condition index methods.

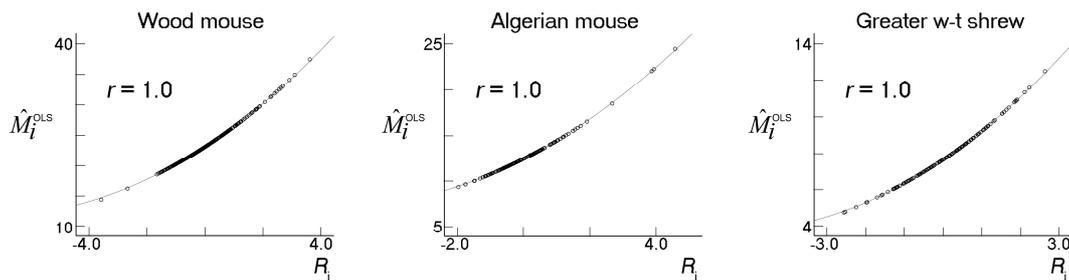
Operationally, some methods generate CIs which are perfectly correlated in linear or exponential fits, and are therefore almost equivalent in terms of their statistical performance (Fig. S7). For example, as both the Residual (R_i) and Relative condition (K_n) indices used the same estimated parameters α and β to compute condition scores from an OLS regression on the whole population, R_i and K_n were perfectly correlated in an exponential fit. Equally, if the Scaled mass index method (\hat{M}_i) had been computed with b_{OLS} from the whole study population instead of b_{SMA} values from the reference group, we would have observed a perfect exponential fit between \hat{M}_i and R_i (see Peig & Green 2009 for mathematical explanation). Finally, if Relative mass (W_r) had employed the same b_{SMA} values as those used by the Scaled mass index, we would have obtained a perfect linear correlation between W_r and \hat{M}_i .

FIG. S7. Statistical equivalence among some condition index methods.

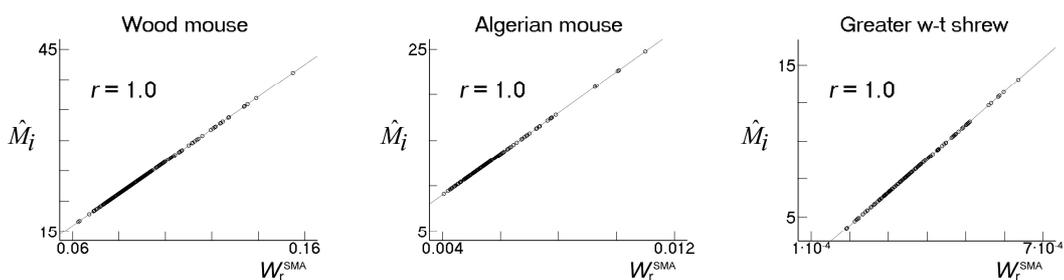
Between Relative condition (K_n) and Residual index (R_i)



Between modified scaled method (\hat{M}_i^{OLS}) and Residual index (R_i) (both using the same b_{OLS} slope)



Between Scaled mass index (\hat{M}_i) and modified relative condition method (W_r^{SMA}) (both using the same b_{SMA} slope)



However, despite the close statistical relationship between some methods, they are not mathematically equivalent (Table S7). The condition scores are different among methods, hence different results may arise when particular CIs are analysed with respect to other variables of interest (e.g. when relating CIs to body components, survival, etc). Importantly, even when perfectly correlated, different methods produce CIs differing in values and units which themselves are of biological importance. According to Packard and Boardman (2009), allometric analyses should be performed on values expressed in the original scale (and units) because it is the scale of interest. For example, focusing on units, providing absolute units of mass (e.g. in g) makes the \hat{M}_i index simple to interpret compared to the BMI or K index which aim to identify individuals that are “relatively heavier or lighter” based on units of kg m^{-2} or kg m^{-3} . Likewise, \hat{M}_i values can be understood more intuitively than CIs provided by the Residual index (R_i) that are

based on differential body mass on a log scale. It is important to note that recovery of the original scale by back-transforming R_i values via antilogs is fraught with problems (Peig & Green 2009). Furthermore, providing comprehensive CIs measures as the whole body mass (in g) for a given length makes the \hat{M}_i particularly suitable for further data description and testing (e.g. the ratios between CIs from two individuals can readily be computed to assess how many times one is heavier than the other). In contrast, this type of description is not straightforward for the R_i index owing to the existence of negative values.

In principle, the lack of negative values could be seen as an advantage for the K_n and W_r indices to facilitate straightforward comparisons of CIs between individuals. However, this is not entirely true since these indices are not symmetrical near 1, that is, the value where the measured mass equals the mass predicted according to these particular methods. For instance, an individual with “good condition” weighing twice the predicted individual mass would produce $K_n = 2$, i.e. one whole unit above the expected ‘normal’ value of one. However, an individual with “bad condition” weighing one half of its predicted mass would have $K_n = 0.5$, i.e., only half a unit below the ‘normal’ value of one. Paradoxically, the arithmetic mean of these two extreme K_n scores is 1.25 (i.e. over one), which illustrates how CI scores above ‘normal’ values are weighted more than those below. Thus, this method can be systematically biased towards “good condition” when further computations are carried out (e.g. note in Table S7 that the arithmetic mean for individual K_n and W_r values were slightly above 1 even though the mean R_i values were zero).

This illustrates a drawback of the implicit assumption made by K_n and W_r that $CI = 1$ represents the “normal” situation. By definition, the Residual index also has the condition that $\sum R_i = 0$, where positive scores are assumed to indicate “good condition” and negative scores “bad condition”. Thus, K_n , W_r and R_i indices provide CI scores which have already been implicitly compared to a condition status presumed to be the “normal” situation (i.e., the average for the study population). In contrast, the Scaled mass index makes no value judgement as to the state of condition of an individual animal, the interpretation of \hat{M}_i scores being necessarily context dependent. Importantly, \hat{M}_i scores are also particularly easy to compare between different studies of the same species.

TABLE S7. Descriptive statistics for condition indices (CIs) and examples of individual values in three small mammal species as computed by different CI methods: the Scaled mass index (\hat{M}_i), Residual index (R_i), Body mass index (BMI), Fulton's index (K), Relative condition (K_n) and Relative mass (W_r).

	\hat{M}_i at $L_0^{(*)}$ (g)	R_i (ln g)	BMI (kg m ⁻²)	K (kg m ⁻³)	K_n (\emptyset)	W_r (\emptyset)
Wood mouse						
(n = 272)						
#1	41.11	3.62	3.76	45.68	1.62	1.72
#2	27.13	1.17	2.80	28.71	1.17	1.21
...
#271	22.31	-0.06	2.40	23.20	0.99	1.01
#272	22.12	-0.09	2.39	22.94	0.99	1.01
$\bar{X} \pm SE$	23.41 \pm 0.22	0.00 \pm 0.06	2.42 \pm 0.02	24.79 \pm 0.26	1.01 \pm 0.01	1.04 \pm 0.01
Min–Max	16.57, 41.11	-3.57, 3.62	1.42, 3.76	17.70, 45.68	0.62, 1.62	0.67, 1.72
Algerian mouse						
(n = 144)						
#1	22.61	2.67	2.86	42.36	1.48	1.84
#2	24.76	3.87	3.43	47.27	1.77	2.05
...
#143	16.57	0.51	2.08	30.99	1.08	1.35
#144	14.05	-0.18	1.88	26.64	0.97	1.16
$\bar{X} \pm SE$	12.73 \pm 0.22	0.00 \pm 0.08	1.97 \pm 0.03	24.86 \pm 0.39	1.01 \pm 0.01	1.07 \pm 0.02
Min, Max	9.11, 24.76	-1.97, 4.57	1.46, 3.87	17.94, 47.27	0.75, 1.96	0.77, 2.05
Greater white-toothed shrew						
(n = 142)						
#1	10.44	0.58	2.16	35.87	1.11	1.09
#2	10.39	2.64	2.47	33.45	1.63	1.46
...
#141	9.9	1.53	2.23	32.73	1.33	1.23
#142	7.61	0.1	1.71	25.15	1.02	0.95
$\bar{X} \pm SE$	7.96 \pm 0.17	0.00 \pm 0.08	1.75 \pm 0.03	26.61 \pm 0.61	1.02 \pm 0.02	0.95 \pm 0.02
Min, Max	4.23, 14.03	-2.56, 2.64	0.98, 2.85	13.37, 48.66	0.62, 1.63	0.56, 1.46

* L_0 = 98.22 mm for Wood mice, L_0 = 79.92 mm for Algerian mice, and L_0 = 67.03 mm for shrews.

References

- Peig, J. & Green, A.J. (2009) New perspectives for estimating body condition from mass/length data: the Scaled mass index as an alternative method. *Oikos*, 118, 1883–1891.
- Packard, G.C. & Boardman, T.J. (2009) A comparison of methods for fitting allometric equations to field metabolic rates of animals. *Journal of Comparative Physiology B*, 179, 175–182.