International Journal of Obesity (1995) 19, 644–652

Statistical considerations regarding the use of ratios to adjust data

DB Allison[1], F Paultre[1], MI Goran[2], ET Poehlman[3] and SB Heymsfield[1]

[1]Obesity Research Center, St Luke's/Roosevelt Hospital Center, Columbia University College of Physicians and Surgeons, 1090 Amsterdam Ave, 14th Floor, New York, NY10025; [2]Department of Nutrition Sciences, Energy Metabolism Research Unit, University of Alabama, Birmingham, Alabama; [3]Baltimore Veterans Affairs Medical Center, University of Maryland School of Medicine, Baltimore, Maryland, USA

OBJECTIVE: The use of ratios to adjust data (i.e. 'index' variables) is common in obesity and related research. The rationale for the use of ratios often seems to be the desire to control or eliminate the influence of the variable in the denominator. The purpose of this paper is to gain a greater appreciation of the statistical assumptions underlying ratios and their impact on data interpretation.

RESULTS: We demonstrate the limitations of the indiscriminant use of ratios to adjust data. Specifically, we show that: (1) given linearity, a zero intercept between the numerator and denominator are necessary and sufficient conditions for a ratio to remove the confounding effects of the denominator; (2) seemingly minor departures from a zero intercept can have major consequences on the ratio's ability to control for the denominator; (3) the ratio of two normally distributed variables cannot be normally distributed, and this may violate the assumptions of subsequent parametric statistical analyses; (4) the use of ratios affects the error distribution of the data which may also violate the assumptions of subsequent parametric statistical analyses; (5) the use of ratios cannot easily take nonlinear effects between the numerator and denominator into account; (6) the use of ratios can introduce spurious correlations among the ratios and other variables; (7) the use of ratios can create interpretive difficulties. We also clarify that the mean of ratios is not necessarily equivalent to the ratio of the means of the numerator and denominator. Finally, we present and discuss formulae for the reliability of ratios and residuals.

CONCLUSION: Because of the above issues, we question the indiscriminant use of ratios and advocate that investigators consider regression-based approaches as alternatives.

Keywords: ratios; index variables; regression; statistical control; data analysis

Introduction

The use of ratios of variables (frequently called 'index' variables)[1] is common in obesity and related research. The motivation for the use of such ratios often seems to be the desire to 'normalize', 'standardize', 'adjust' or 'control' for the variable in the denominator.[2] Some ratios commonly used in obesity research include waist-to-hip ratio (WHR),[3] body mass index (BMI),[4] percent fat,[5] resting metabolic rate (RMR) divided by fat free mass (FFM),[6,7] the activity factor (TEE/RMR; where TEE is total energy expenditure),[54] total cholesterol to HDL cholesterol,[8] and VO_2 max divided by body weight.[10] For further listing of specific ratios and discussion of their performance see Kronmal[2] and Tanner.[9]

Although several investigators have pointed out specific problems with ratios,[6,10,11] a more general and comprehensive treatment has not been offered to the obesity research community. In our own work we have often used ratios (e.g. BMI, subscapular to triceps skinfold ratio)[12] and have at times been unsure of the appropriateness of this practice. Thus, our purpose here is to examine the statistical assumptions underlying ratio approaches and possible consequences associated with their use and misuse.

Ratios of variables can be used for several purposes.[2] From our reviews of the obesity related literature, it appears that one can distinguish between the ratio of two variables being used simply to predict or estimate a third variable and the ratio being used in the context of 'causal modeling' or 'hypothesis testing'. In the former case, one is simply seeking a combination the two component variables that provides the optimal prediction or estimation of a third variable. Although the ratio may meet this end, there are an infinite number of other ways to combine two pieces of information which may prove better. This paper *is not* primarily focused on this issue (see Draper and Smith[13] and references therein for more on this topic).

This paper is primarily concerned with the use of ratios in the context of 'causal modeling' or 'hypothesis testing'. Within this context we again distinguish between two uses of ratios. In the first case, investigators are implicitly stating that some 'special' aspect of the ratio has a relation with another variable that is not accounted for by either the numerator or denominator, individually or additively. For example, an investigator might posit that it is the combination of one's waist circumference and hip circumference that influences morbidity. This is equivalent to saying that it is this particular *interaction* of waist (W) and hip (H) that influences morbidity.[2] When W/H is expressed as $W*H^{-1}$, the problem becomes familiar as one of moderated multiple regression.[14,15,16] The generally accepted way to test this sort of interaction hypothesis is to simultaneously regress the dependent variables (e.g. some measure of morbidity) on the two main effects (e.g. W and H^{-1}) and the product

Correspondence: DB Allison.

term (e.g. W^*H^{-1}). If the regression coefficient for the product term is significantly different from zero then one can claim that the multiplicative combination W^*H^{-1} has some influence on morbidity above and beyond the effects of W or H^{-1} or their additive combination.[2] In other words, such an analysis would show that there is something 'special' about the ratio of waist to hip. For further discussion of these topics see.[14,15,16,17]

The final use of ratios, i.e. 'controlling for the denominator', will be the major focus of this paper. For example, investigators have questioned whether obese persons have lower RMR's than lean persons. In absolute terms, obese persons tend to have *higher* RMR's than lean persons due in part to their greater fat free mass (FFM).[6] This information leads to a reformulation of the question of whether obese persons have lower RMR's than lean persons *after* accounting for their greater lean body mass. In an effort to address this question, several investigators compared lean and obese persons using the ratio of RMR to FFM. In other words, RMR is divided by FFM to control for FFM.

In regard to controlling via ratios the concerns are twofold. First, when one's theoretical model suggests the need to control for a random variable (X), then one should *fully* control for X. Second, if one uses ratios as variables in parametric statistical analyses, the ratios should meet the assumptions of these analyses. The use of a particular ratio is justifiable to the extent that it satisfies these criteria. Alternatives should be considered when a particular ratio does not meet these criteria. In the following nine sections we review issues influencing the extent to which ratios are likely to meet these criteria, discuss selected statistical properties of ratios, and offer guidelines for their interpretation and use. Finally, we briefly discuss the general regression approaches which offer flexible alternatives to the use of ratios to control for variables.

Specific issues

The necessity of a zero intercept

Several authors [9,10,18,19] have pointed out that using a ratio is a legitimate method to control for a denominator only when the sample intercept (b_0) of the regression of the numerator (Y) on the denominator (X) is zero. However, to our knowledge no formal and exact proof has been offered for this statement. The central issue hinges on the desire for an index that is independent of (i.e. uncorrelated with) the denominator. If Y is the variable of interest and one attempts to control for another variable, say X, by forming a ratio, (i.e. Y/X), then one's success in statistically controlling for X is measured by how close $r_{Y/X,X}$, the correlation between Y/X and X, is to zero. In the interests of space, these proofs are not presented here but can be obtained by writing to the first author. In these proofs, we show that, under linearity, when $b_0 = 0$ the *expected value* of $r_{Y/X,X}$ is zero. Furthermore, we show that under bivariate normality, the population intercept (β_0) equaling zero is a necessary and sufficient condition for the population correlation ($\rho_{Y/X,X}$) to equal zero.

Consequences of seemingly minor departures from a zero intercept

In reality, b_0 is a *continuous* random variable. The probability of a *continuous* random variable assuming any particular value is zero. In other words, the probability of b_0 being *exactly* zero is zero (even when $\beta_0 = 0$). This raises the question, 'how close to zero does b_0 need to be for the ratio to perform acceptably?' Unfortunately, this is a question with no easy answer. The answer will depend on the criteria for acceptable performance.

One rule that has been suggested is to use the ratio Y/X only when b_0 is *not* significantly different from zero.[21] However, it is possible for $r_{Y/X,X}$ to be *significantly different* from zero even when b_0 is *not significantly different* from zero. This possibility is of concern because, assuming that the purpose of a ratio is to control for the denominator, a nonsignificant correlation between the ratio and its denominator is really a minimum criterion for acceptable performance. This combination will occur whenever these two inequalities hold:

$$r_{Y/X,X}\sqrt{\frac{n-2}{1-r^2_{\frac{Y}{X},X}}} > T_{crit} \tag{1}$$

$$\frac{b_0}{s_{Y/X}\sqrt{\dfrac{1}{n}+\dfrac{\bar{X}^2}{(n-1)s_x^2}}} < T_{crit} \tag{2}$$

where T_{crit} is the value of t that cuts off the upper $\alpha/2\%$ of the t distribution with n-2 degrees of freedom. The easiest way to prove this possibility is by illustration. We provide a hypothetical sample data set in Table 1 along with statistical calculations. As demonstrated, b_0 is *not* significant but $r_{Y/X,X}$ is significant, indicating that taking the ratio Y/X does not adequately control for X. Given this example we can see that testing whether b_0 is significantly different from zero is a questionable heuristic for deciding when b_0 is too far from zero to support using a ratio. We think our example is even more striking given that X and Y are themselves perfectly uncorrelated. Thus, in this example, divid-

Table 1 A hypothetical data set

	Y	*X*	*Y/X*
	1.86	1.91	0.98
	3.56	4.15	0.86
	2.36	2.02	1.17
	1.45	3.52	0.41
	2.26	1.94	1.17
	3.39	1.03	3.30
	3.35	3.53	0.95
	1.45	2.20	0.66
	0.49	2.93	0.17
	1.72	3.59	0.48
Mean	2.19	2.68	1.02
s.d.	0.95	0.95	0.82

$b_0 = 2.19$ ($t = 2.179$; $df = 8$; P = 0.061).
$r_{y/x,x} = -0.675$ (P = 0.032).

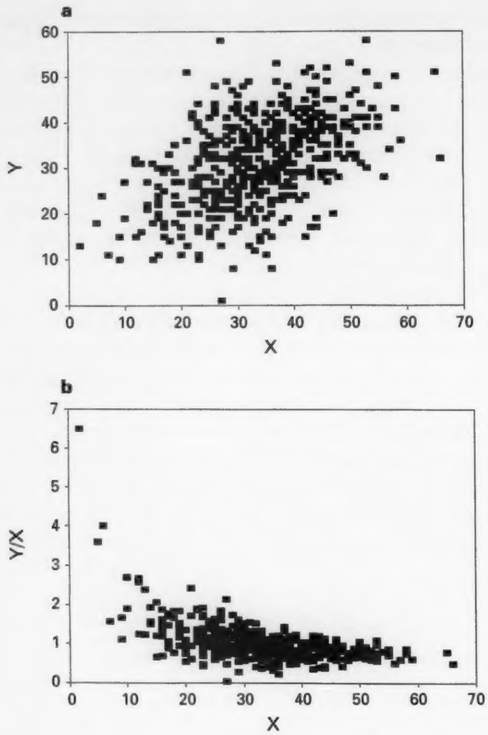

Figure 1a Simulated homoscedastic data. (n = 500; ρ = 0.5). **b** Heteroscedastic data. (Data from **a** after ratio is taken).

ing Y by X not only fails to control for X, but it actually induces a significant covariation with X.

An alternative heuristic might be to test the significance of b_0 at some very lenient level (e.g. $\alpha = 0.20$) and test all substantive hypotheses at some stringent levels (e.g. $\alpha = 0.05$ or 0.01). This should markedly reduce the probability of residual confounding[20] due to a non-zero b_0 introducing a statistically significant degree of confounding into the relationship of Y/X with Z, where Z is some third variable of interest correlated with X. Residual confounding occurs when one believes one has controlled for a confounder but the confounder's influence has not been completely removed from the variable system under study.[20] However, we have not rigorously evaluated this heuristic and the quality of its performance remains speculative. Moreover, such heuristics may be unnecessary because, as we discuss in the final section, the use of ratios can be incorporated in regression approaches that do not depend on a zero intercept.

Taking a ratio affects the error distribution

Homoscedasticity (homogeneity of variance) is a basic assumption of regression analysis and other parametric techniques.[22] Homoscedasticity occurs when the conditional variance of Y is constant for all values of X. In contrast, heteroscedasticity occurs when the conditional variance of Y is not constant for all values of X. Figures 1a and 1b illustrate homoscedasticity and heteroscedasticity, respectively.

Unfortunately, if Y is homoscedastic when regressed on X, the ratio Y/X will be heteroscedastic with respect to X and, in all probability, to other variables correlated with X.[2] Lynn and Bond[21] explain that Y can always be expressed as a polynomial function of X, i.e.,

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \cdots + \beta_{n-1} X_i^{n-1} + e_i \qquad (3)$$

where n is the number of observations (subjects). Thus, the ratio, Y/X, is equivalent to:

$$\frac{Y_i}{X_i} = \frac{\beta_0}{X_i} + \beta_1 + \beta_2 X_i + \cdots + \beta_{n-1} X_i^{n-2} + \frac{e_i}{X_i} \qquad (4)$$

Of interest is the final term of equation 4, e_i/X_i. By definition e_i, the error term, has mean zero, variance σ_e^2, and covariance with X equal to zero for all i's. This implies that the expected value of (e_i/X_i given X_i) is also zero but its variance varies as a function of X (specifically, ($\sigma^2_{e/Xi}$ given X_i) = σ_e^2/X_i^2).

This heteroscedasticity that is potentially introduced by the ratio transformation violates the assumptions of virtually all subsequent parametric statistical analysis of the ratio Y/X when it is considered relative to X or to variables correlated with X. These violations can either increase the type I error rate or increase the type II error rate depending on the specific pattern.[23] In other words, *P*-values and significance tests can no longer be trusted when heteroscedasticity is present.

Our main point here is not to state that taking a ratio will always produce heteroscedasticity, but rather to demonstrate that taking a ratio does affect the error distribution. In some cases these effects can even be desirable. In fact, Bowerman and O'Connell[24] discuss how ratios can often be used to restore homoscedasticity. However, having been alerted to the fact that taking ratios alters the error distribution, the case for carefully checking the residuals for homoscedasticity becomes even more compelling.

Finally, significance tests and *P*-values from parametric statistical analyses on heteroscedastic data can be particularly inaccurate when group n's are unequal or other assumptions (e.g. normality) are also violated. This brings us to our next point.

The distribution of ratios

A perusal of any issue of *IJO* indicates that most investigators rely on parametric statistical inferences. It is well known that many parametric statistical tests are robust to violations of normality.[25] However, some statistical analyses, such as the path analyses and structural equation modeling frequently used in the quantitative genetic analyses of obesity (e.g. References 26,27), may be more sensitive to violations of normality.[28,29] Moreover, violations of normality can be of greater concern when sample sizes are small and/or variances are heterogenous[23] and we showed in the previous section that variance of ratios may be heterogenous. Thus, it is important to try to minimize departures from normality whenever possible.

It is therefore noteworthy that the ratio of two normally distributed variables cannot be normally distributed.[30] Under some circumstances, the departure from normality can be quite extreme. Unfortunately, there is no *simple* mathematical expression for the distribution of a ratio of two variables or its moments even if the two components of the ratio are normally distributed (methods for deriving the distribution and its moments are available.[31,32,33,34] Thus, it would be prudent for investigators who are considering using ratios to evaluate the observed frequency distribution for normality (tests for normality can be found in D'Agostino, *et al*.[35] This is not to say that it is not also wise to test the distributions of variables used in parametric analyses even if they are not ratios. If the ratio markedly deviates from normality, either the ratio should not be used or, if possible, it should be transformed to approximate normality.[36] Software to implement tests of normality and select appropriate transformations is available.[55] Again, if marked non-normality occurs with small samples, unequal samples, or variance heterogeneity, then the type I or type II error rates may increase.

Ratios cannot easily accommodate nonlinearity

Thus far, we have assumed that the relationship between Y and X is linear. Under this restriction, we showed that taking the ratio Y/X adequately controls for X when $\beta_0 = 0$. That is, when the model is:

$$Y_i = \beta_1 X_i + e_i \tag{5}$$

In this case, both a linear regression of Y on X and taking the ratio Y/X will adequately control for X.

However, it is well known that relationships among biological and psychological variables are often nonlinear, being characterized by curved rather than straight lines.[37] One possible such curve is characterized by the quadratic equation below:

$$Y_i = \beta_1 X_i^2 + e_i \tag{6}$$

If this is the 'correct' model relating Y to X, then neither a linear regression of Y on X nor the ratio Y/X will adequately control for X. However, both methods can be easily adapted to correct for non-linearity. In the ratio method, one can simply divide Y by X^2 to remove any confounding effects of X. Perhaps the most well known example of this method is Body Mass Index or Quetelet's Index, weight divided by height squared (Kg/m^2).[38] In the regression approach we can simply regress Y on X^2.

More generally, any time the model can be characterized by the following form:

$$Y_i = \beta_1 X_i^k + e_i \tag{7}$$

either the regression approach or the ratio approach can easily accommodate the data.

The regression approach simply regresses Y on X^k while the ratio approach simply divides Y by X^k. An example of the latter is Bazette's[39] formula for the corrected QT interval from an electrocardiogram ($QT_c = QT/(RR)^{1/2}$). Here Y is the QT interval, X is RR (hear rate) and k = 0.5.

The challenge in using the approach dictated by equation[7] is knowing the value of k. It is possible to derive k on the sample data through nonlinear regression. However, k must be solved for iteratively[40] and use of this approach becomes more complex. Moreover, there may be no value of k that will provide an adequate characterization of the data.

However, if we switch to a model that is nonlinear in the dependent variables but linear in its parameters, then the process becomes simpler. As stated previously, any bivariate data set can be fully characterized in the following manner:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \cdots + \beta_{n-1} X_i^{n-1} + e_i \tag{8}$$

Equation 8 is nonlinear in the independent variables and can, therefore, characterize curved data. However, because it is linear in its parameters it can be rapidly solved by ordinary least squares with standard regression software. This use of multiple regression with polynomial terms is a flexible approach that can *completely* remove any relationship between Y and X, linear or nonlinear. The same cannot be said of the ratio approach. Even if $\beta_0 = 0$, under the model in equation 8, the ratio approach can only be used if one and only one β_j is not zero, where j = 1 to n–1. Thus, in terms of accounting for non-linearity, the ratio approach is a far less flexible approach to statistical control than is the use of regression based approaches.

At this point we should point out that if, instead of equation 7, one believed the data were characterized by a multiplicative model such as:

$$Y = \beta_0 X^k e \tag{9}$$

then a ratio approach might be quite desirable and appropriate. In this case, one need not necessarily solve for the parameters iteratively. Rather, by taking the natural logarithms of both sides of the equation, the equation becomes one that is easily solved in closed form by ordinary least squares regression software. Once the parameters are solved, the resulting equation is exponentiated and returned to its original form. This useful and general approach is customarily referred to as *allometric scaling*.[62-68] While this approach is seemingly useful and can be quite useful, it should be noted that it makes a number of assumptions especially about the error distribution, introduces biases into the obtained coefficients that require correction, and has a number of other nuances.[62-68] A complete discussion of allometric scaling is beyond the scope of this paper. The references cited provide more details.

Spurious correlations

Not only can a ratio fail to completely control for its denominator, but it can actually introduce confounding and, therefore, spurious correlations.[2,14] This is the main point of

Pearson's classic paper.[41] Pearson showed that under certain not too restrictive assumptions (i.e., multivariate normality and means that are large relative to their standard deviations) the correlation between two ratios can be derived directly from the correlations among the components, their means, and their standard deviations. When both Y and X are divided by the same denominator (e.g. z) this yields the following:

$$r_{\frac{Y}{Z},\frac{X}{Z}} \approx \frac{r_{Y,X}V_Y V_X - r_{Y,Z}V_Y V_Z - r_{X,Z}V_X V_Z + V_Z^2}{\sqrt{V_Y^2 + V_Z^2 - 2r_{Y,Z}V_Y V_Z}\sqrt{V_X^2 + V_Z^2 - 2r_{X,Z}V_X V_Z}} \quad (10)$$

where $r_{Y/Z,X/Z}$ is the sample correlation between Y/Z and X/Z, and V_j is the coefficient of variation of j (i.e. $V_j = \sigma_j/\mu_j$).

Pearson[41] points out that if Y, X and Z are completely uncorrelated (i.e. each observation is just a random triplet) and the three coefficients of variation (i.e. standard deviations divided by the means) are equal, then the correlation of Y/Z with X/Z will be 0.5.

Consider the following hypothetical example. Following the work of Drewnowski,[42] an investigator wishes to test whether preference for fat is correlated with preference for sugar among free living humans. The investigator measures food intake via seven-day food diaries and calculates daily caloric intake. The means, standard deviations and correlations for this hypothetical data are displayed in Table 2. Although our data are hypothetical, concern with the use of ratios in these situations is not.[43,56]

As can be seen, the correlation between sugar intake and fat intake is 0.15, suggesting a small positive association. However, the investigator may wish to control for total caloric intake because people who eat more in general will almost certainly eat more fat and sugar (note the positive correlations). If the investigator calculated the partial correlation of fat and sugar intake after controlling for total caloric intake, the answer would be 0.00. In other words, after controlling for total caloric intake there is no correlation between sugar preference (operationalized by sugar intake) and fat preference (operationalized by fat intake). However, if the investigator chose to 'control for' total caloric intake by correlating percent of total calories from fat with percent of total calories from sugar, a markedly different picture would emerge. Substituting the appropriate values in Table 2 into equation 9, we obtain the following result: $r_{\%\ \text{fat},\ \%\ \text{sugar}} = 0.86$. Thus, in this example, attempting to control by the ratio method actually increases the confounding due to the denominator and leads to an extremely biased and erroneous result.

Table 2 Hypothetical summary statistics for food intake data

	Fat calories (F)	*Sugar calories (S)*	*Total calories (T)*
Mean	520	1040	2600
s.d.	80	150	700

$r_{F,S} = 0.15$
$r_{F,T} = 0.30$
$r_{S,T} = 0.50$

It should be pointed out that the preceding explanation and example do not imply that correlating two ratios with a common denominator always leads to spurious correlation, but rather that it can lead to spurious correlation. Finally, it is important to define what we mean by a spurious correlation. In the example above, when we say that the observed correlation of 0.86 between percent of kcal from fat and percent kcal from sugar is spurious, we do not mean to imply that the correlation is not really 0.86. On the contrary, in our hypothetical example the correlation between %fat and %sugar really *is* 0.86. However, it is spurious in that it is *not* the correlation of sugar and fat intake after controlling for total caloric intake.

The issue of controlling for total caloric intake in epidemiological studies has received considerable attention[43,56-59] and the current zeitgeist in epidemiological studies is to use residuals instead of ratios.[56] Although this position is not without controversy,[57-59] most of the issues revolve not around the repudiation of ratios but rather on the interpretation and testing of regression coefficients when residuals are employed.[57-60] In addition to examining the aforementioned literature, investigators contemplating the use of residuals might wish to consult other sources on this topic (e.g. References 52, 60, 61).

Interpretive difficulties

As several authors (e.g. References 1, 44–46) have previously pointed out, the use of ratios can at times lead to interpretive difficulties (for example see the accompanying paper by Goran *et al.*).[5] In part, these difficulties stem from condensing two variables into one.

Consider an investigator who reports that among a group of subjects (group A) the RMR/FFM ratio is greater than among some other group of subjects (group B). This finding can be interpreted in one of four ways:

(1) members of group A have higher RMR's than members of group B;
(2) members of group A have lower values of FFM;
(3) members of group A have higher values of both RMR and FFM but the between group difference in RMR is proportionately greater than the between group difference in FFM, or finally;
(4) members of group A have lower values of both RMR and FFM but the between group difference in FFM is proportionately greater than the between group difference in RMR.

With knowledge of only the ratios, any one of these four potential conclusions is possible. Thus, if investigators do rely on ratios of variables, it is helpful if they also report means and variances of the numerator and denominator variables so the various possibilities above can be disentangled. In addition, reporting the correlation between the numerator and denominator is helpful. This correlation should be reported within groups if groups are being compared because, as we explain in the next section, the mean of a ratio is not only a function of the numerator and denominator means and variances but also their intercorrelation.

The ratio of means is not necessarily the mean ratio
One minor point which is not always appreciated is that the ratio of means is not necessarily equal to the mean ratio. Specifically, the ratio of means is defined as:

$$\frac{\bar{Y}}{\bar{X}} = \frac{\frac{1}{n}\sum Y_i}{\frac{1}{n}\sum X_i} = \frac{\sum Y_i}{\sum X_i} \tag{11}$$

In contrast, the mean ratio is defined as:

$$\text{Mean of } \left(\frac{Y_i}{X_i}\right) = \frac{1}{n}\sum\frac{Y_i}{X_i} = \frac{1}{n}\sum\frac{b_0 + b_1 X_i + e_i}{X_i} \tag{12}$$

Substituting $(\bar{Y}-b_0)/(\bar{X})$ for b_1 yields:

$$\frac{1}{n}\sum\frac{Y_i}{X_i} = \frac{\bar{Y}}{\bar{X}} + b_0\left[\left(\frac{1}{n}\sum\frac{1}{X_i}\right) - \frac{1}{\bar{X}}\right] + \frac{1}{n}\sum\frac{e_i}{X_i} \tag{13}$$

It can be seen that equation 13 is *not* equivalent to equation 11. Thus if one wishes to calculate the mean ratio, one will typically need the raw data to do so. Although the mean ratio can be roughly estimated, it cannot be calculated simply from the two components' means as several authors seem to have done (e.g. References 47,48). However, if more detailed summary statistics are available (i.e. s_x, s_y, $\bar{X}$, $\bar{Y}$, $r_{X,Y}$), Pearson[41] showed that the mean ratio can be approximated by:

$$\frac{1}{n}\sum\frac{Y_i}{X_i} \approx \frac{\bar{Y}}{\bar{X}}\left(1 + \frac{s_X^2}{(\bar{X})^2} - r_{X,Y}\frac{s_Y s_X}{\bar{Y}\bar{X}}\right) \tag{14}$$

Inspection of equation 14 reveals that the ratio of the means will only be equal to the mean of the ratios, in this approximation, when the following condition is met:

$$\frac{s_X \bar{Y}}{s_y \bar{X}} = r_{X,Y} \tag{15}$$

Equation 16 is met only when the intercept (b_0) of the regression of the numerator on the denominator is zero. Alternatively, using the exact formula in equation 13 we find that the *expected value* of the mean ratio equals the ratio of the means only when $b_0 = 0$.

As an example consider data from the recently published Zutphen study[49] some of which is reproduced in Table 3. The investigators correctly calculated the subscapular to triceps skinfold ratio (STR) for a cohort of middle aged men. Subsequently, some of the men developed gallstones and others did not. The mean difference in STR between the two groups was 0.1. Given that the difference in STR seemed to confer increased risk for gallstones, we might speculate that a difference in STR of 0.1 is clinically significant. Thus, it is interesting to note that *if* the mean of the ratio had been incorrectly calculated as the ratio of the mean of the numerator and the mean of the denominator, the obtained results would have been off by 0.1.

Table 3 Zutphen study data (Moerman *et al.*)

	With gallstones	*Without gallstones*
Variable	$n = 806$	$n = 54$
Tricipital skinfold (mm)[2]	9.8	10.1
Subscapular skinfold (mm)[3]	14.7	16.5
STR: Subscapular-to-triceps skinfold thickness ratio (mm)	1.6	1.7

Reliability of ratios and residuals
Unfortunately, data are virtually always measured with error. Therefore, consideration of the reliability of measurement becomes a crucial aspect of any study. A major alternative to the use of ratios is linear regression analysis and/or the calculation of residuals. Thus, it may also be instructive to consider the relative reliabilities of ratios and residuals.

An approximate formula for the reliability of a ratio was first derived by Cronbach[50] based on Pearson's[41] approximation to the variance of a ratio. Because we have shown in previous sections that the ratio approach is only a viable strategy when $\beta_0 = 0$, we restrict our comparisons to that situation. If $\beta_0 = 0$ then the formula for the reliability of the ratio reduces to:

$$\rho_{\frac{Y}{X}\frac{Y}{X}} \approx \frac{\rho_{Y,Y} + \rho_{Y,X}^2(\rho_{X,X} - 2)}{1 - \rho_{Y,X}^2} \tag{16}$$

Malgady and Colon-Malgady[51] independently derived a formula for the reliability of a residual. The reliability of the residual ($\rho_{e,e}$) turns out to be equal to the approximation of $\rho_{Y/X,\ Y/X}$ when $\beta_0 = 0$. Therefore, under the conditions where a ratio may be acceptable (i.e. when $\beta_0 = 0$), the ratio and residual are equally reliable.

When using or contemplating the use of ratios or residuals, investigators may wish to calculate the reliability of these derived scores using equation 15. However, the results may be disconcerting. In general, if the components are measured with 'good' reliability, the derived scores will

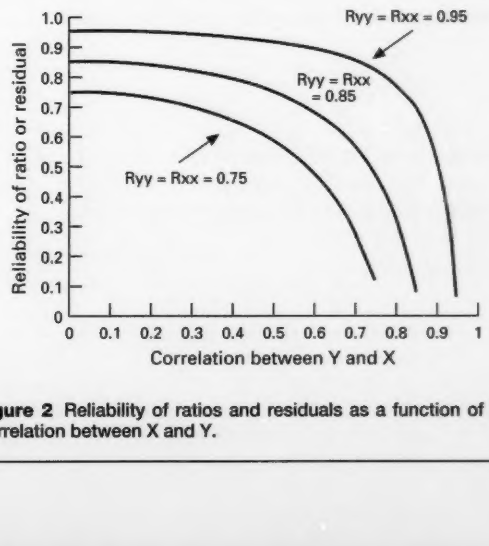

Figure 2 Reliability of ratios and residuals as a function of the correlation between X and Y.

have only 'moderate' reliability and components with only 'moderate' reliability will generally yield derived scores with 'poor' reliability. This is illustrated in Figure 2.

General commentary

Thus far we have primarily described properties of ratios and highlighted potential concerns. Here, we offer some general commentary followed by some guidelines regarding the control of variables.

First, it is important to limit the scope of our comments. Our paper is not meant to apply to all ratios in all contexts. Rather, we restricted our focus to the use of ratios of random variables in the context of controlling for one of the variables (i.e. the denominator). Thus, our comments do not apply to all quantities arrived at by the process of division (e.g. risk ratios, odds ratios, F-ratios, likelihood ratios, etc.).

Second, all of our examples have included variables which are scaled such that they have a fixed zero point. This is typical of biological data. However, this is *not* true of most psychological test scores. Because the zero point is not fixed, scores on psychological tests can generally be scaled to have any mean and variance. In this case, one's choice of scaling will dramatically alter the results obtained with ratios. This is not the case with regression analyses, whose inferential results are unaffected by linear transformations. Thus, because psychological data have no fixed zero point (i.e. interval or ordinal data), the use of ratios is strongly discouraged.

Regarding guidelines, put quite simply, the optimal approach to the data is one that gets the job done (i.e. fully controls for the appropriate variable(s)) and meets the assumptions of the statistical analyses employed. As we have shown, the only way to be sure that the ratio Y/X fully controls for X is for the intercept (b_0) of the regression of Y on X to be zero. Because b_0 will never be exactly zero, the simple ratio approach will always be questionable.

One way to overcome this difficulty might be to subtract b_0 from each Y forming the variable $Y' = Y - b_0$. The ratio Y'/X will now fully control for X because the intercept of the regression of Y' on X will be exactly zero. One could now implement a more appropriate approach to the use of ratios as a means of control by proceeding in three steps.

In the first step, compute the regression equation:

$$Y_i = b_0 + b_1 X_i + e_i \tag{17}$$

In the second step, calculate Y_i' as $Y_i - b_0$. In the third step, take the ratio Y_i'/X_i and regress it on the other variable of interest (e.g. z) yielding the regression equation:

$$\frac{Y_i'}{X_i} = a_0 + a_1 Z_i + e_i \tag{18}$$

Alternatively, one could use the new ratio as an independent variable in the following equation

$$Z_i = a_0 + a_1 \frac{Y_i'}{X_i} + e_i \tag{19}$$

At this point the reader may recall that the more conventional regression equation:

$$Y_i = b_0 + b_1 Z_i + b_2 X_i + e_i \tag{20}$$

also fully controls for any linear effects of X and can also control for nonlinear effects through incorporation of polynomials [In an effort to control for X prior to regressing Y on Z, some investigators take the residual of equation 17 and regress it on Z. It is worth noting that this is not always formally equivalent to computing the full regression model (equation 20) and testing the significance of the regression coefficient for Z. For further discussion see (References 52, 53, 61).]

What then is the difference between these approaches? The major difference resides in the distribution of residuals. One of these equations will probably meet the assumptions of regression analysis (i.e. that the residuals are normally and independently distributed with mean zero and constant variance) better than the other. Residual plots and other more sophisticated methods (see any good regression text for details) can be used to evaluate the extent to which the residuals satisfy the regression assumptions. Thus in choosing between these approaches, the better approach is the one that more closely satisfies the regression assumptions.

Conclusion

We hope that the preceding exposition proves helpful to investigators attempting to control for selected variables in analyses and contemplating the use of ratios. We believe that the simple approach of dividing Y by X to control for X will often prove inadequate. Similarly, a simple linear approach to controlling for X can also prove inadequate at times. Every method has assumptions that one must test and limitations that one must be aware of. Although we have tried to offer practical guidelines, we must emphasize that there is no substitute for thoroughly exploring one's data, trying alternative approaches, and testing the assumptions of one's chosen technique.

Acknowledgements
This work was supported in part by National Institutes of Health grants DKY2618 & DK26687. David Allison is supported by an NIDDK FIRST award (R29DK47256-01A1). Furcy Paultre is supported by a NIH predoctoral fellowship.

We gratefully acknowledge Dr J J McArdle for his helpful discussion on this paper and Rachel L Weber, Melissa I Kezis and Orlando A Velazco for their assistance in preparing the manuscript.

References

1 Kerlinger FN, Foundations of behavioral research: *Educational and Psychological Inquiry*. Holt, Rinehart and Winston, Inc: New York, 1964, pp. 616–618
2 Kronmal RA. Spurious correlation and the fallacy of the ratio standard revisited. *J Roy Stat Soc A* 1992; **156 (part 3)**: 379–392.
3 Krotkiewski M, Björntorp P, Sjostrom L, Smith U. Impact of obesity on metabolism in men and women-importance of regional adipose tissue distribution. *J Clin Invest* 1983; **72**: 1150–1162.
4 Garrow JS, Webster J. Quetelet's index (W/H^2) as a measure of fatness. *Internat J Obes* 1985; **9**: 147–153.
5 Goran MI, Allison DB, Poehlman ET (submitted): Issues relating to normalization of body fat content. *Int J Obes* 1995; **19**: 638–643.
6 Ravussin E, Bogardus C. Relationship of genetics, age, and physical fitness to daily energy expenditure and fuel utilization. *Am J Clin Nutr* 1989; **49**: 968–975.
7 Platte P *et al.* Resting metabolic rate and total energy expenditure in acute and weight recovered patients with anorexia nervosa and in healthy young women. *Int J Eating Disorder* 1994; **16**: 45–52.
8 Stampfer MJ *et al.* A prospective study of cholesterol, apolipoproteins, and the risk of myocardial infarction. *New England J Med* 1991; **325**: 373–381.
9 Tanner JM. Fallacy of per-weight and per-surface area standards, and their relation to spurious correlation. *J Appl Physiol* 1949; **2**: 1–15.
10 Toth MJ. Examination of data normalization procedures for expressing peak Vo_2 data. *J Appl Physiol* 1993; **75 (5)**: 2288–2292.
11 Poehlman ET, Toth MJ (submitted). Mathematical ratios lead to spurious conclusions regarding age and gender related differences in resting metabolic rate.
12 Allison DB, Heshka S, Heymsfield SB. Evidence of a major gene with pleiotropic action for a cardiovascular disease risk syndrome in children younger than 14 years. *Am J Dis Child* 1993; **147**: 1298–1302.
13 Draper NR, Smith H. *Applied Regression Analysis.* Wiley: New York, 1981.
14 Cohen J, Cohen P (1983): *Applied Multiple Regression/ Correlation Analysis for the Behavioral Sciences*. Lawrence Erlbaum: Hillsdale, NJ, pp. 73–76.
15 Cohen J. Partialed products *are* interactions; Partialed powers *are* curve components. *Psychol Bull* 1978; **85**: 858–866.
16 Aiken LS, West SG. *Multiple Regression: Testing and Interpreting Interactions*. Sage: Newbury Park.
17 Jaccard J, Turrisi R, Wan CK. *Interaction Effects in Multiple Regression*. Sage: Newbury Park.
18 Katch V. Correlation *vs* ratio adjustments of body weight in exercise-oxygen studies. *Ergonomics* 1972; **15 (No. 6)**: 671–680.
19 Katch VR. Use of the oxygen/body weight ratio in correlational analyses: spurious correlations and statistical considerations. *Med Sci Sports* 1973; **5 (No. 4)**: 253–257.
20 Becher H. The concept of residual confounding in regression models and some applications. *Stat in Med* 1992; **11**: 1747–1758.
21 Lynn M, Bond CF Jr. Conceptual meaning and spuriousness in ratio correlations: The cases of restaurant tipping. *J App Soc Psyc* 1992; **22 (No. 4)**: 327–341.
22 Berry WD. *Understanding Regression Assumptions*. Sage: Newbury Park, 1992, pp.67–81.
23 Alexander RA, Govern DM. A new and simpler approximation for ANOVA under variance heterogeneity. *J Ed Stat* 1994; **19 (No. 2)**: 91–101.
24 Bowerman BL, O'Connell RT. *Linear Statistical Models*. 2nd edn., Boston, PWS-KENT: 1990, pp. 644–662.
25 Boneau CA. The effects of violations of assumptions underlying the *t*-test. *Psychol Bull* 1960; **57**: 49–64.
26 Allison DB, Neale MC, Heshka S, Heymsfield SB. Race effects in the genetics of adolescents' body mass index. *Int J Obesity* 1994; **18**: 363–368.
27 Korkeila M, Kaprio J, Rissanen A, Koskenvuo M. Effects of gender and age on the heritability of body mass index. *Int J Obesity* 1991; **15**: 647–654.
28 Bentler PM, Stein JA. Structural equation models in medical research. *Stat Methods in Med Res* 1992; **1**: 159–181.
29 Bollen KA. *Structural Equations with Latent Variables.* Wiley-Interscience: New York, 1989, pp. 416–418.
30 Sokal RR, Rohlf FJ. *Biometry*. WH Freeman: New York, 1980, pp. 17–19.
31 Hinkley DV. On the ratio of two correlated normal random variables. *Biometrika* 1969; **56 (3)**: 635–639.
32 Fieller EC. The distribution of the index in a normal bivariate population. *Biometrika* 1932; **24**: 428–440.
33 Johnson NL, Kotz S. *Distributions in Statistics: Continuous Multivariate Distributions.* John Wiley & Sons, Inc: New York, 1976, pp. 122–124.
34 Marsaglia G. Ratios of normal variables and ratios of sums of uniform variables. *J Am Stat Assoc* 1965; **60**: 193–204.
35 D'Agostino RB, Belanger A. D'Agostino RB Jr. A suggestion for using powerful and informative tests of normality. *Am Statist* 1990; **44**: 316–321.
36 Box GEP, Cox DR. An analysis of transformations. *J Roy Stat Soc, Series B* 1964; **26**: 211–252.
37 Klieber M. Body size and metabolic rate. *Physiol Rev* 1947; **27**: 511–539.
38 Quetelet LAJ. A treatise on man and the development of his faculties. *Obesity Res* 1994; **2**: 72–85.
39 Bazette HC. An analysis of the time relation of electrocardiograms. *Heart* 1920; **7**: 353.
40 Chambers J. Fitting nonlinear models: Numerical techniques. *Biometrika* 1973; **60**: 1–13.
41 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 Roy Soc London* 1987; **60**: 489–497.
42 Drewnowski A. Sensory preferences for fat and sugar in adolescence and adult life. In: *Nutrition and Chemical Series in Ageing*. *Ann NY Acad Sci* 1989; **561**: 243–250.
43 Martin-Moreno JM. Adjustment for total caloric intake in nutritional studies: an epidemiological perspective. *Eur J Clin Nutr* 1993; **47 (suppl. 2)**: S51–S52.
44 Darlington RB. *Regression and Linear Models.* McGraw-Hill: New York, 1990, pp. 77–79.
45 Zolman JF. *Biostatistics: Experimental Design and Statistical Inference.* Oxford University Press: New York, 1993, p. 95.
46 Johnson AG. *Statistics.* RK Merton (ed.) Harcourt Brace Jovanovich: New York, 1988, pp. 45–48.
47 Garrow JS. Is it possible to prevent obesity? *Infusionstherapie* 1990; **17**: 28–31.
48 Zillikens MC, Conway JM. Anthropometry in blacks: Applicability of generalized skinfold equations and differences in fat patterning between blacks and whites. *Am J Clin Nutr* 1990; **52**: 45–51.
49 Moerman CJ, Berns MPH, Smeets WM, Kromhout D. Regional fat distribution as risk factor for clinically diagnosed gallstones in middle-aged men: a 24-year follow-up study (the Zutphen Study). *Int J Obesity* 1994; **18**: 435–439.
50 Cronbach LJ. The reliability of ratio scores. *Ed Psych Meas* 1941; **1**: 269–278.
51 Malgady RG, Colon-Malgady G. Comparing the reliability of difference scores and residuals in analysis of covariance. *Ed Psych Meas* 1991; **51**: 803–807.
52 Maxwell SE, Delaney HD, Manhiemer JM. ANOVA of residuals and ANCOVA: correcting an illusion by using model comparisons and graphs. *J Ed Stat* 1985; **10 (No. 3)**: 197–209.

53 Bollen KA, Ward S. Ratio variables in aggregate data analysis: Their uses, problems, and alternatives. *Sociol Methods Res* 1979; **7 (No. 4)**: 431–450.
54 World Health Organization: Energy and protein requirements. *Tech Rep Ser*, 1985.
55 Allison DB, Gorman BS, Kucera EM (in press). UNICORN: A computer program for transforming data to normality. *Educat Psychol Measurement*.
56 Willett W, Stampfer MJ. Total energy intake: implications for epidemiologic analyses. *Am J Epidemiology* 1986; **124**: 17–27.
57 Pike MC, Peters RK, Bernstein L. RE: Total energy intake: Implications for epidemiological analyses. *Am J Epidemiol* 1993; **137**: 811–813.
58 Palmgren J, Kushi LH. RE: Total energy intake: Implications for epidemiological analyses. *Am J Epidemiology* 1991; **133**: 1291–1293.
59 Howe GR, Mkiller AB, Jain M. RE: Total energy intake: Implications for epidemiological analyses. *Am J Epidemiology* 1976; **124**: 157–159.
60 Cutter HSG, O'Farrell TJO. Alternatives to regression residuals as measures of change after the intake of alcohol. *Alcoholism: Clin Exp Res* 1987; **11**: 186.
61 van den Burg W, Lewis C. Testing multivariate partial, semipartial and bipartial correlation coefficients. *Multivariate Beh Res* 1990; **25**: 335–340.
62 Nevill AM, Ramsbottom R, Williams C. Scaling physiological measurements for individuals of different body size. *Eur J Appl Physiol* 1992; **65**: 110–117.
63 Sprugel DG. Correcting for bias in log-transformed allometric equations. *Ecology* 1983; **64**: 209–210.
64 Martin RD, Barbour AD. Aspects of line-fitting in bivariate allometric analyses. *Folia Primatol* 1989; **53**: 65–81.
65 Hartman SE. A critique of some regression adjustments used in allometric size 'correction' in numerical taxonomy. *Am J Phys Anthropol* 1983; **62**: 305–309.
66 Smith RJ. Logarithmic transformation bias in allometry. *Am J Phys Anthropol* 1993; **90**: 215–228.
67 Reiss MJ. *The Allometry of Growth and Reproduction.* Cambridge University Press: New York, 1989.
68 Mosimann JE, James FC. New statistical methods for allometry with application to Florida red-winged blackbirds. *Evolution* 1979; **33**: 444–459.