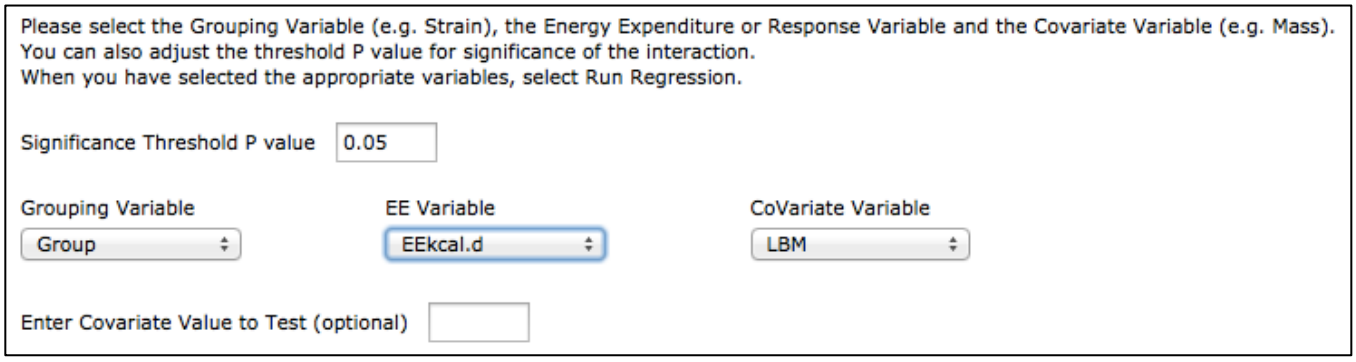
Module 3: Using the MMPC Web Tool to Revisit a "Hypometabolic" Phenotype in Myostatin Knockout Mice

Module 1 used data obtained in myostatin deficient and control mice to illustrate how a mathematical artifact can confound group comparisons based on traditional ratio EE normalization by forcing the ratio to be correlated with the body mass compartment. Traditional ratio normalization creates a strong negative correlation between EE normalized as (EE /LBM) and LBM. In our example, the myostatin KO mice have greater LBM than control mice, which leads to the ratio-normalized EE being significantly hypometabolic in myostatin KO mice in comparison to control mice based on a t-test. This module provides a "hands on" vignette analyzing the myostatin data to demonstrate the use of the MMPC ANCOVA tool to compare EE in a way that appropriately adjusts for differences in body mass.

About the dataset The data in MyostatinKOdata.csv were measured in the Nutrition and Obesity Research Center's Energy Balance Core at the University of Washington, and represent a subset of the data published in [1]. Groups are KO, denoting myostatin knockout mice (n=13) and CON, denoting control mice (n=31). CON mice include wildtype mice and mice that were heterozygous for the myostatin mutation. EE is in units of $kcal/d$ and mass variables are in g. Abbreviations are: TBM, total body mass; LBM, lean body mass; FM, fat mass; EEkcal.d, average 24h EE in kcal/d; EE.lbm, ratio of EE to LBM in units of $kcal/d$ per g LBM; EE.tbm, ratio of EE to TBM in $kcal/d$ per g TBM.

Analysis of average 24h EE with lean body mass as the covariate Enter the variables to be analyzed as shown in Figure 1 using the MMPC analysis web site (http://www.mmpc.org/shared/regression.aspx). Leave the Significance Threshold P value at 0.05 for the test to determine whether the slope of EE on LBM differs significantly by group.

Figure 1



Click Run Regression to obtain the output shown in Figure 2.

Figure 2

Program= MMPC Multiple Linear Regression Output					
Datafile=MyostatinKOdata.csv					
Response Variable=EEkcal.d					
Covariate=LBM					
Grouping Variable=Group					
The interaction LBM:Group is NOT significant. (p=0.2991)					
44 total cases. CON = 31, KO = 13					
R Squared = 68.673		Adjusted R Square = 67.1448			
Residual Standard Error 0.6203 with 41 degrees of freedom					
Source	Sum of Squares	df	MeanSquare	F-ratio	P value
Regression	34.5872	2	17.2936	44.9387	4.63807E-11
Residual	15.7779	41	0.384826		

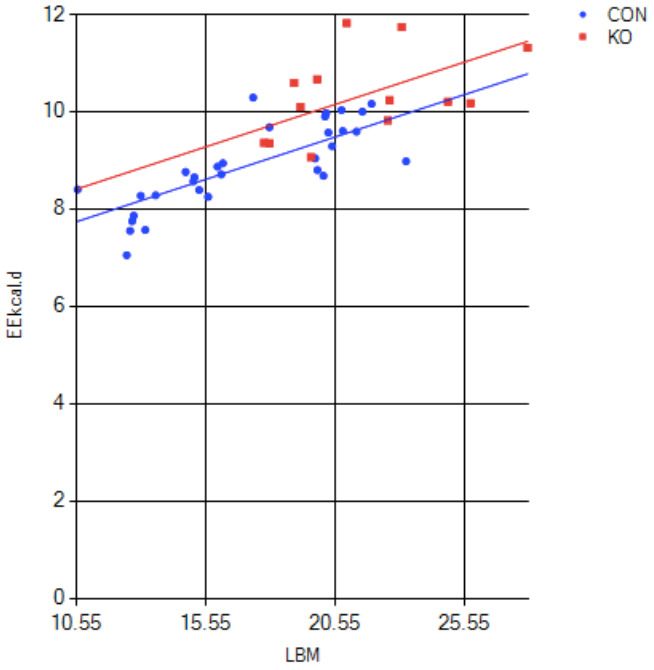
Variables	Estimate	StdError	P Value
(Intercept)	5.9026	0.48737	3.9968E-15
LBM	0.174014	0.0276791	1.69057E-07
Group	0.668386	0.240108	0.00809288

Basic Statistics - Avg (StDev)

Groups	CON	KO
EEkcal.d	8.89 (0.838)	10.34 (0.842)
LBM	17.14 (3.506)	21.66 (3.053)

Model-based Statistics - Avg EEkcal.d (StErr)

	Groups		
Used: LBM	CON	KO	P value
Overall Mean	9.12 (0.117)	9.79 (0.193)	0.008093
Group Mean	8.89 (0.111)	10.34 (0.172)	1.207E-08
Residual Variance	0.27406	0.67322	



Interpretation:
Topmost information box:

- Includes basic information on variable selection and sample sizes.
- The LBM by group interaction term is denoted as LBM:Group.
- The statistical significance of the LBM:Group interaction, shown by the p-value of p=0.2991, indicates that the slope of EE on LBM does not differ between groups at the designated significance threshold value for the interaction term. Accordingly, the regression model excludes the interaction term.
- The R Squared value indicates that the ANCOVA accounts for 68.7% of the total variance in 24h EE (p=4.64 E-11; shown in the line labeled Regression in the table). An explanation of Adjusted R Square of 67.14 is beyond the scope of this vignette.

Variables information box and graph of 24h EE vs. LBM by group:

- Each regression line in the graph is the ANCOVA predicted mean EE given a particular value of LBM. The vertical distance that separates the lines (i.e. as the lines for each group have the same slope, the vertical distance between the lines is equal to the difference in the y-intercepts) equals the group difference, and at any given value of LBM, KO mice have higher predicted mean EE than CON mice. This difference is significant as indicated below.
- The general form of the ANCOVA regression model obtained in this analysis is:
  $Estimated\ group\ mean\ 24h\ EE\ at\ a\ specific\ value\ of\ body\ mass = intercept + estimate\ for\ mass\ \times\ mass\ value + estimate\ for\ group\ if\ group = KO, 0\ if\ group = CON$

  EE group comparisons are typically adjusted to the sample-wide mean of the covariate(s) in the analysis. Here, the sample-wide mean LBM is 18.476 g.

  $Estimated\ mean\ 24h\ EE\ at\ 18.5\ g = 5.902 + 0.174\ \times\ 18.476 + 0.668\ if\ group = KO, 0\ if\ group = CON$
  =9.78 kcal/d for KO, 9.12 kcal/d for CON.

  Note that in this example, the estimate for group in the Variables information box equals the KO minus CON group. In some analyses, the estimate for group may instead be computed as the control group minus experimental group difference. The order of the computation depends on the names used for the groups and the way the WEB site organizes the groups for analysis. Note that because the regression lines in the graph show the mathematical ANCOVA model in visual form, one can tell at a glance how to interpret the estimate for group.
  The values associated with the model are given in the form 'parameter (also called coefficient) estimate ± standard error of the estimate' and are interpreted as follows:

Variables	Estimate	StdError	P Value	
(Intercept)	5.9026	0.48737	3.9968E-15	y intercept (±SEM) for CON group
LBM	0.174014	0.0276791	1.69057E-07	Slope (±SEM) for both groups as the interaction is not significant
Group	0.668386	0.240108	0.00809288	Difference in y-intercept (±SEM) between CON and KO

- On average, 24h EE changes (i.e. slope of line) by 0.174 ± 0.0276 kcal/day per g change in LBM (p=1.69 E-7 = 0.000000169).
- The groups differ in 24h EE by 0.6684 ± 0.2401 kcal/day (p=0.008).

Basic Statistics Box

- This shows basic (unadjusted) means and standard deviations for 24h EE and LBM by group.

Model Based Statistics box

- The line labeled Overall Mean shows the ANCOVA-adjusted group means, standard errors and significance of the group comparison when the comparison is adjusted to mean LBM taken over both groups (i.e., the grand mean for LBM). Stated differently, the Overall Mean information gives the estimated EEs and standard errors that the groups would have if every animal had the same LBM, and this equaled the mean LBM of the entire sample of N=31+13=44 mice. Recall from Module 1 that we are typically interested in answering two questions: (1) To estimate the

group difference in EE (± SE) that would pertain if the groups had the same mean LBM; and, (2) To estimate the mean EE (± SE) that would pertain for each group if the groups had the same mean LBM. To address question 1 the vertical difference between the two lines (which equals the group difference in y-intercepts) is given in the prior summary data (0.6684±0.2401 kcal/d) and the difference is significant (p=0.008). Regarding question 2, for most users the Overall Mean line provides the most important information in the Model Based Statistics box. At equivalent body weights KO have a higher EE (9.79±0.19 kcal/d) EE than WT (9.12±0.12 kcal/d) and the difference is significant at p=0.008.

- However, some users may be interested in an adjusted group comparison that does retain the effect of the group mass difference on the group EE difference. The line labeled Group Mean shows adjusted means, standard errors and the p-value for the group comparison when the comparison is adjusted to the individual group means of LBM. Thus for the CON group, EE for every animal is adjusted to a LBM of 17.14 g, while for the KO group, EE values are all adjusted to a LBM of 21.66 g (see Basic Statistics box), and the group comparison is performed at these values. This procedure yields that same group difference as obtained in the numerator in the equation for a Student's independent samples t-test. The advantage to the adjusted approach is that the variance of the group EE estimates and the variance of the EE difference between groups are reduced by accounting for variance in EE due to variance in body mass. This yields a more precise estimate of the group difference in an analysis that is intended to embody both the effect on EE of a group difference in body mass and a mass-independent EE component. Combining the information in the Group Mean line with the information provided in the Overall Mean line can help an investigator estimate how much of a group EE difference is independent of body mass (usually the major question), and how much is due to differences in body mass (which is of obvious importance to defining a group's holistic EE phenotype).
- What can you conclude from this analysis? After accounting for the average effect on EE of LBM per se, the KO mice have a higher EE than control mice. As EE is increased, this finding could help explain the relatively lean phenotype of the KO mice.

Sensitivity analysis A diagnostic called 'dfbeta' enables users to identify individual animals that might exert excessive influence on the analysis. Consider the dfbeta plot for group in Figure 3: The y-axis shows the change in the estimated group difference when a particular animal, denoted by Index, is deleted from the analysis. Index 1 corresponds to the first row of data, index 2 corresponds to the second row, and so forth.

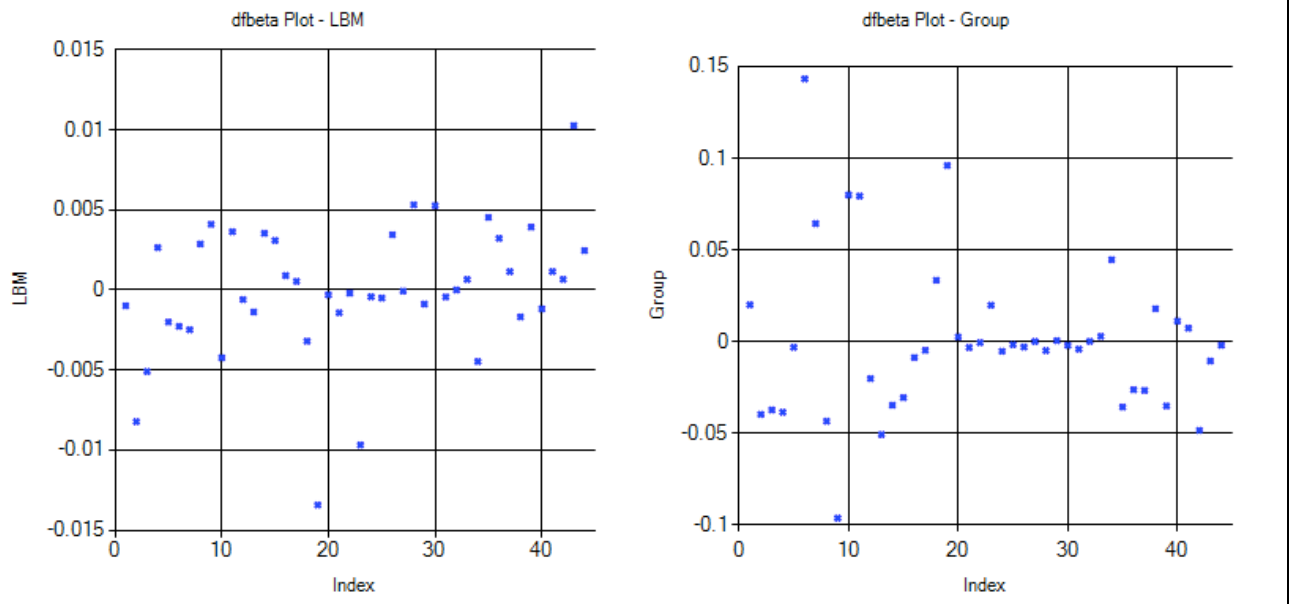
Removing index 6 from the analysis results in the largest change in the estimated group difference (about 0.14). More precisely, if index 6's data are removed from the analysis and the regression is re-run on the remaining cases, the estimated group difference changes by 0.143 EE units because:
(group difference estimate with full dataset = 0.668) – (group difference estimate without index6 = 0.525) = 0.143.

Thus, one way to compute the value of the group difference coefficient when an index is excluded is to *subtract* the index's dfbeta from the original group difference estimate.

However, one can almost instantly obtain not just the revised group difference, but also the revised standard error and significance calculations. Simply position the cursor over a dfbeta point to see the values. For the most extreme dfbeta (corresponding to index 6), the index-deleted group difference is 0.525, the standard error is 0.226 and the revised p-value is 0.0255.

Potentially influential cases are easy to note by identifying dfbeta values that are large relative to the magnitude of the coefficient estimate. One should be concerned with any single observation that drastically changes the results of the analyses (which does not happen in the present example).

Figure 3.



ANCOVA-adjusted EE is uncorrelated with LBM In module 1 we stated that key requirement of valid normalization is that the adjusted EE values must be minimally or un-correlated with the body mass compartment in question (assuming that the question is to estimate EE "as if" both groups had the same mass values). Figure 4 demonstrates this and compares ANCOVA to ratio normalized EE.

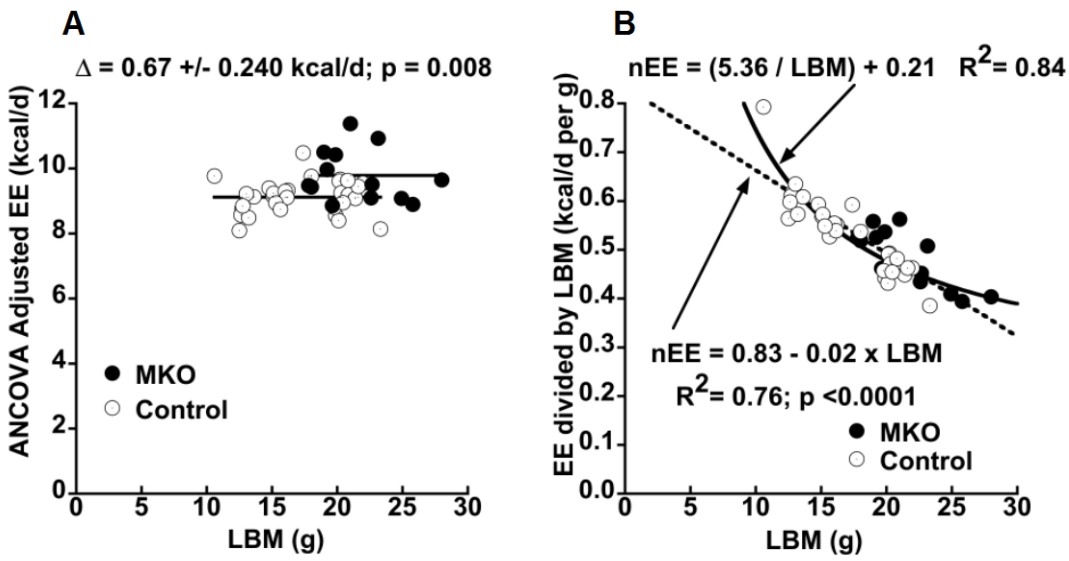


Figure 4. A: The average slope of the ANCOVA adjusted EE vs. body mass is zero, i.e., the adjusted EE values are uncorrelated with LBM. **B**: By contrast, ratio normalized EE (nEE) is strongly correlated with LBM.

Acknowledgements: This module was written by Karl J. Kaiyala[1], Charles F. Spiekerman[2], Robert H. Podolsky[3] and Owen P. McGuinness[4]. The MMPC ANCOVA web tool was created by Richard McIndoe[5].

[1]School of Dentistry, Department of Oral Health Sciences, University of Washington, Seattle, WA; [2]School of Public Health and Department of Biostatistics, University of Washington, Seattle, WA; [3]Department of Family Medicine and Public Health Sciences, Wayne State University School of Medicine, Detroit MI; [4]Department of Molecular Physiology and Biophysics, Vanderbilt University School of Medicine, Nashville, TN; [5]Coordinating and Bioinformatics Unit, Georgia Regents University, Augusta, GA.

The authors were supported by funding from the National Institute of Health (KJK: U24DK076169, P30DK035816; CFS: U24DK076169; RHP: R03-DA033453; OPM: DK59637, DK20593, DK078188, DK043748; RM: U24DK076169) and the NIDDK Mouse Metabolic Phenotyping Center Program.

References

1. Choi SJ, Yablonka-Reuveni Z, Kaiyala KJ, Ogimoto K, Schwartz MW, Wisse BE (2011) Increased energy expenditure and leptin sensitivity account for low fat mass in myostatin-deficient mice. Am J Physiol Endocrinol Metab 300: E1031-1037.