HEB report with analysis code

Dillon Niederhut

September 25, 2014

```
## Warning: NAs introduced by coercion
```

Descriptives

```
##      logTEE           logW             logRMR           PAL       
##  Min.   :1.423   Min.   :-1.3143   Min.   :1.188   Min.   :0.5816  
##  1st Qu.:2.381   1st Qu.: 0.2467   1st Qu.:1.937   1st Qu.:1.3628  
##  Median :2.845   Median : 0.8976   Median :2.305   Median :1.9666  
##  Mean   :2.818   Mean   : 0.7831   Mean   :2.323   Mean   :2.6038  
##  3rd Qu.:3.358   3rd Qu.: 1.7288   3rd Qu.:2.723   3rd Qu.:3.4501  
##  Max.   :3.686   Max.   : 2.2109   Max.   :3.290   Max.   :6.7933  
##  NA's   :35                        NA's   :52      NA's   :87
```

```
##
## Allenopithecus       Alouatta          Aotus         Ateles     Callicebus 
##              1              3              1              1              1 
##      Callimico     Callithrix          Cebus     Cercocebus  Cercopithecus 
##              6              1              4              6              2 
##        Eulemur        Gorilla          Homo      Hylobates          Lemur 
##             12              3             53              4             12 
##      Lepilemur         Macaca       Marmoset     Microcebus            Pan 
##              1             16              3             18             16 
##          Papio          Pongo       Saguinas        Saimiri   Symphalangus 
##             11              9              2              1              1 
##        Tamarin 
##              7
```

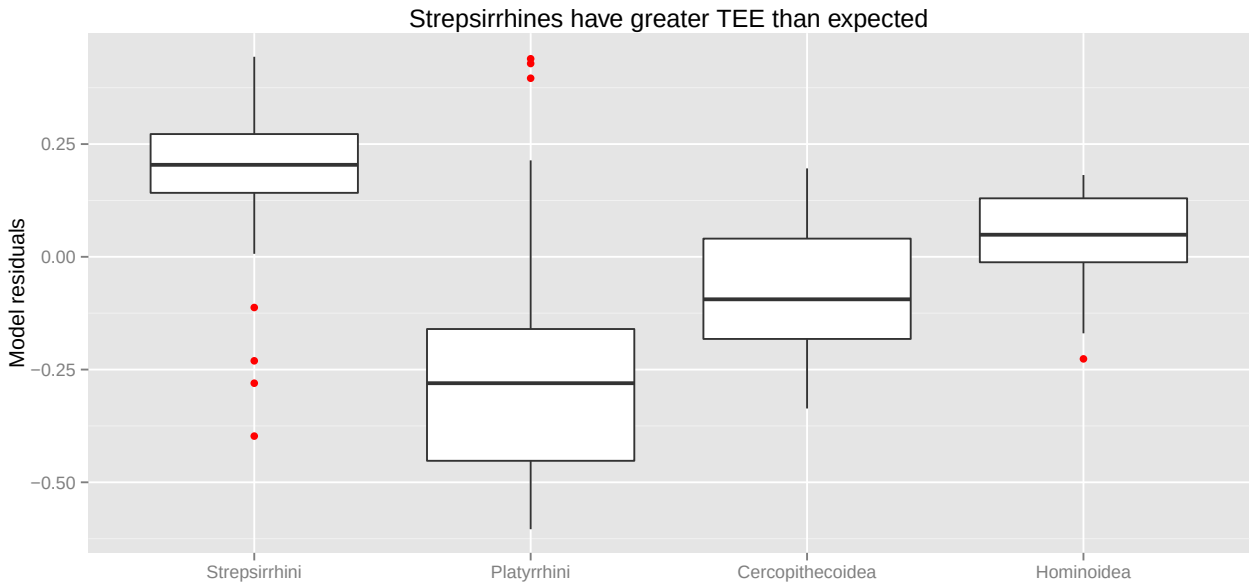
```
##
##                                        CAN 2003 
##                                              23 
##                                     Coehlo 1986 
##                                              33 
##                              Leonard & Robertson 1997 
##                                              29 
##                           Leonard Galloway Ivakine 1997 
##                                               4 
##                      Leonard Katzmarzyk Stephen Ross 1995 
##                                               4 
##                                  Pontzer et al 2010 
##                                               3 
##                                  Pontzer et al 2014 
##                                              15 
##                          Power Tardiff Power Layne 2003 
##                                               6 
##                        Rothman Dierenfeld Hintz Pell 2008 
##                                               1
```

```
## Rothman Dierenfeld Hintz Pell 2008, Smith Jungers 1997
##                                                     1
##                                    Schmid Speakman 2000
##                                                    18
##                                       Simmen et al 2010
##                                                    23
##                                Westerterp & Speakman 2008
##                                                    35

##
## Doubly Labeled Water            Factorial           Heart Rate
##                   94                   54                    8
## Respiratory Quotient
##                   39
```

Split happens with haplorhines

Differences in energy use between primate clades is obvious

```
##
## Call:
## lm(formula = resid ~ family, data = data.primates)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.57330 -0.07886  0.01085  0.09641  0.66818
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.02393    0.02227  -1.075 0.285064
## family1     -0.19970    0.03206  -6.229 1.04e-08 ***
## family2      0.13918    0.03559   3.911 0.000165 ***
## family3      0.05837    0.03655   1.597 0.113320
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2131 on 103 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.3932, Adjusted R-squared:  0.3755
## F-statistic: 22.25 on 3 and 103 DF,  p-value: 3.479e-11

## Warning: Removed 35 rows containing non-finite values (stat_boxplot).
```




```
## Warning: Removed 35 rows containing non-finite values (stat_boxplot).
```

Humans are largest contributor to ape variability

```
##
## Call:
## lm(formula = logTEE ~ logW + pan, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6200 -0.1178  0.0358  0.1730  0.4978
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.43403    0.02323 104.801   <2e-16 ***
## logW         0.50774    0.01795  28.290   <2e-16 ***
## pan1        -0.04933    0.09973  -0.495    0.622
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2376 on 157 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.8377, Adjusted R-squared:  0.8357
## F-statistic: 405.3 on 2 and 157 DF,  p-value: < 2.2e-16

##
## Call:
## lm(formula = logTEE ~ logW + gorilla, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.61915 -0.12290  0.03626  0.17399  0.49770
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.43308    0.02316 105.055   <2e-16 ***
## logW         0.50686    0.01809  28.020   <2e-16 ***
## gorilla1    -0.01169    0.14079  -0.083    0.934
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2378 on 157 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.8375, Adjusted R-squared:  0.8354
## F-statistic: 404.6 on 2 and 157 DF,  p-value: < 2.2e-16

##
## Call:
## lm(formula = logTEE ~ logW + pongo, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.61989 -0.12514  0.03193  0.17217  0.50041
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.43403    0.02312 105.274   <2e-16 ***
## logW         0.50990    0.01812  28.140   <2e-16 ***
## pongo1      -0.09293    0.10069  -0.923    0.357
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2372 on 157 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.8384, Adjusted R-squared:  0.8363
## F-statistic: 407.2 on 2 and 157 DF,  p-value: < 2.2e-16

##
## Call:
## lm(formula = logTEE ~ logW + hylobates, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.62097 -0.12183  0.03491  0.17218  0.49562
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.43488    0.02309 105.439   <2e-16 ***
## logW         0.50662    0.01772  28.586   <2e-16 ***
## hylobates1  -0.29508    0.23741  -1.243    0.216
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2367 on 157 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.8391, Adjusted R-squared:  0.837
## F-statistic: 409.3 on 2 and 157 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = logTEE ~ logW + human, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.60407 -0.11874  0.01286  0.16050  0.44336
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.41366    0.02278 105.976  < 2e-16 ***
## logW         0.44542    0.02358  18.893  < 2e-16 ***
## humanHuman   0.19878    0.05288   3.759  0.00024 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2278 on 157 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.8509, Adjusted R-squared:  0.849
## F-statistic:   448 on 2 and 157 DF,  p-value: < 2.2e-16

## Start:  AIC=-458.61
## logTEE ~ logW
##
##           Df Sum of Sq    RSS     AIC
## + human    1   0.73320 8.1473 -470.40
## <none>                 8.8805 -458.61
## + pongo    1   0.04792 8.8326 -457.48
## + gorilla  1   0.00039 8.8801 -456.62
##
## Step:  AIC=-470.4
## logTEE ~ logW + human
##
##           Df Sum of Sq    RSS     AIC
## <none>                 8.1473 -470.40
## + gorilla  1  0.064935 8.0824 -469.68
## + pongo    1  0.015237 8.1321 -468.70

##
## Call:
## lm(formula = logTEE ~ logW + human, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.60407 -0.11874  0.01286  0.16050  0.44336
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.41366    0.02278 105.976  < 2e-16 ***
## logW         0.44542    0.02358  18.893  < 2e-16 ***
## humanHuman   0.19878    0.05288   3.759  0.00024 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

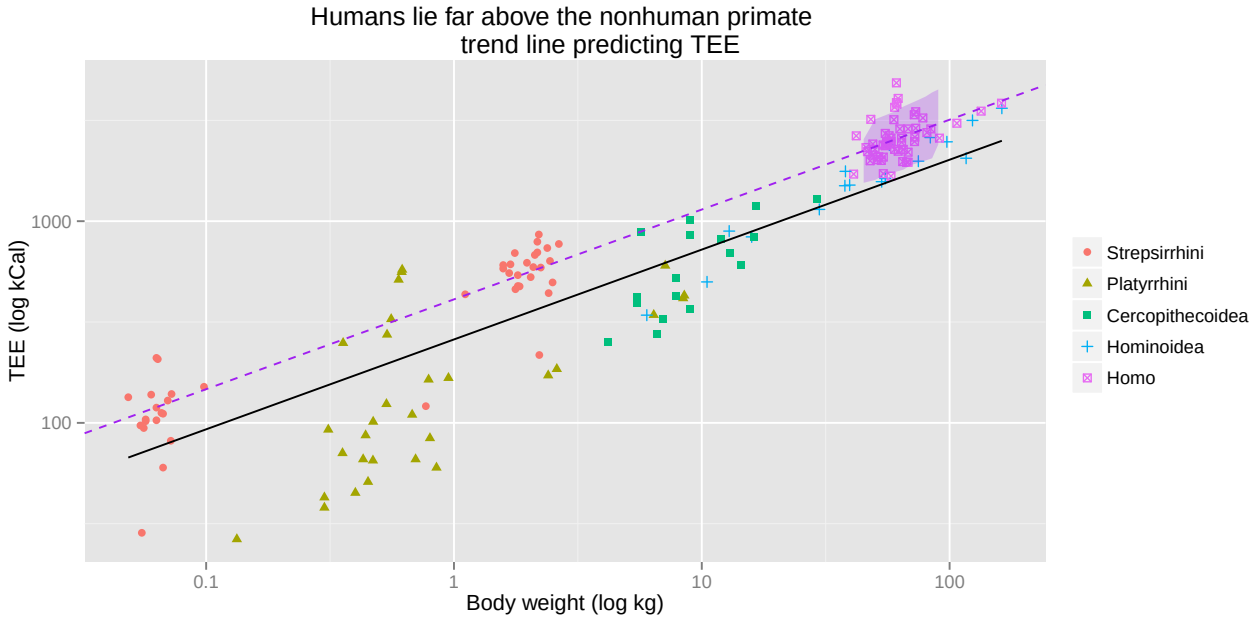
```
## Residual standard error: 0.2278 on 157 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.8509, Adjusted R-squared:  0.849
## F-statistic:   448 on 2 and 157 DF,  p-value: < 2.2e-16
```

Total energy expenditure is higher than predicted for a primate

- Dotted line is human trend line, shaded area is WHO established range of human TEE variation

```
## Warning: Removed 35 rows containing missing values (stat_smooth).
```

```
## Warning: Removed 35 rows containing missing values (geom_point).
```




```
## Warning: Removed 35 rows containing missing values (stat_smooth).
```

```
## Warning: Removed 35 rows containing missing values (geom_point).
```

A model that only uses body weight is

```
## [1] 0.002828195
```

as likely to minimize information lost as one that includes a variable for humans
Delta r-squared is

```
## [1] 0.01341711
```

, a significantly better prediction at F =

```
## [1] 14.12884
```

and p =

```
## [1] 3.369076e-08
```

A Fisher's Exact Test reveals that the probability of having all human data points above the trend line is:

```
##
##  Fisher's Exact Test for Count Data
##
## data:
## p-value = 1.525e-05
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.08416755 0.46500960
## sample estimates:
## odds ratio
##  0.2081821
```

This is not because humans have higher RMR

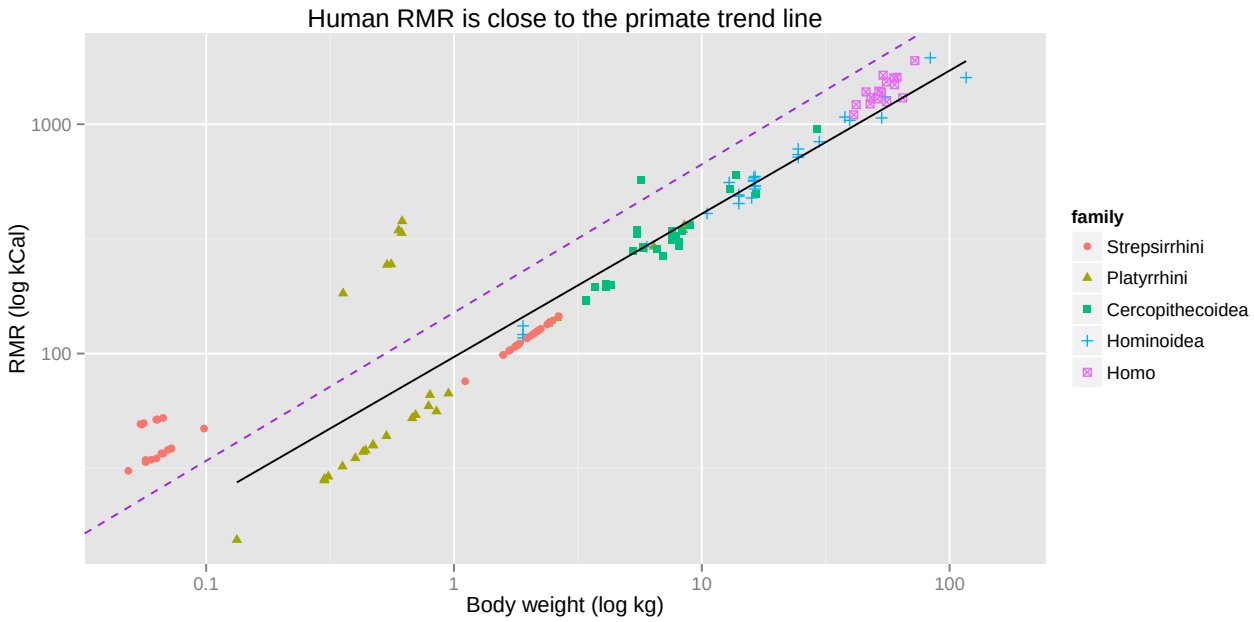
Human RMR is greater than expected for a primate of our body size

Streps removed from dataset for RMR calculations

```
## Start:  AIC=-344.94
## logRMR ~ logW
##
##         Df Sum of Sq    RSS     AIC
## <none>               3.3335 -344.94
## + human  1  0.064548 3.2690 -344.93
```

```
## Warning: Removed 13 rows containing missing values (stat_smooth).
```

```
## Warning: Removed 52 rows containing missing values (geom_point).
```

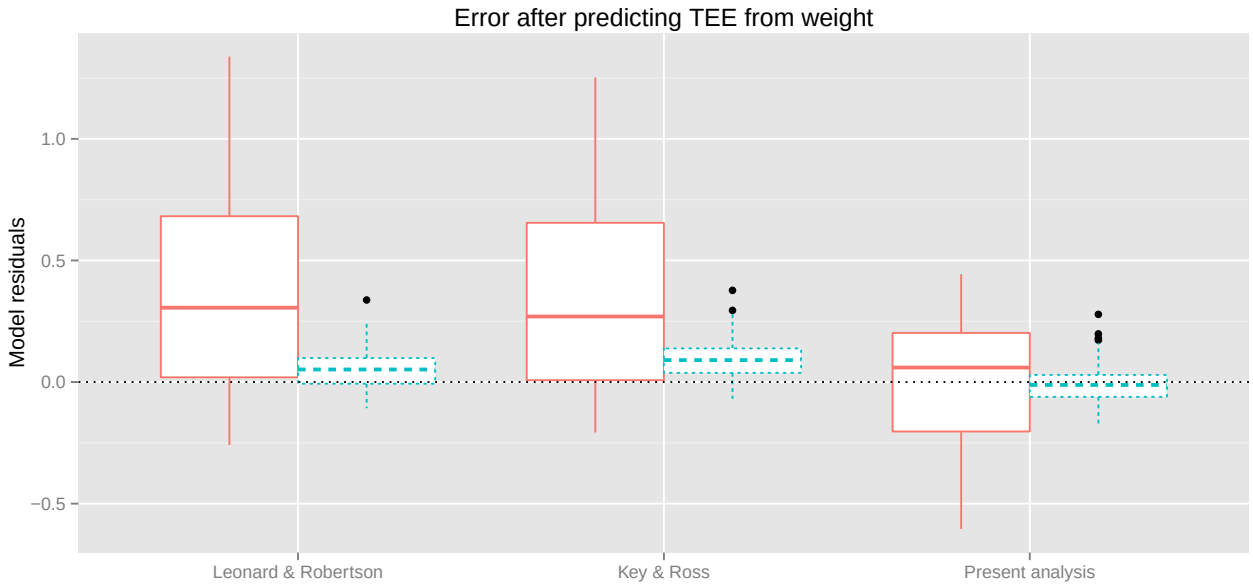



```
## Warning: Removed 13 rows containing missing values (stat_smooth).

## Warning: Removed 52 rows containing missing values (geom_point).
```

Human TEE is still significantly higher if you subtract the extra RMR value from humans

```
## Start:  AIC=-344.94
## logRMR ~ logW
##
##         Df Sum of Sq    RSS     AIC
## <none>               3.3335 -344.94
## + human  1  0.064548 3.2690 -344.93

##
## Call:
## lm(formula = removedTEE ~ logW + human, data = data.nostrep)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.38128 -0.09379 -0.01411  0.05368  0.68231
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.20127    0.02840  77.523  < 2e-16 ***
## logW         0.59286    0.02684  22.092  < 2e-16 ***
## humanHuman   0.16151    0.04571   3.533 0.000594 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1809 on 114 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.9078, Adjusted R-squared:  0.9062
## F-statistic: 561.1 on 2 and 114 DF,  p-value: < 2.2e-16
```

This model provides a better fit for both humans and nonhumans than previous analyses

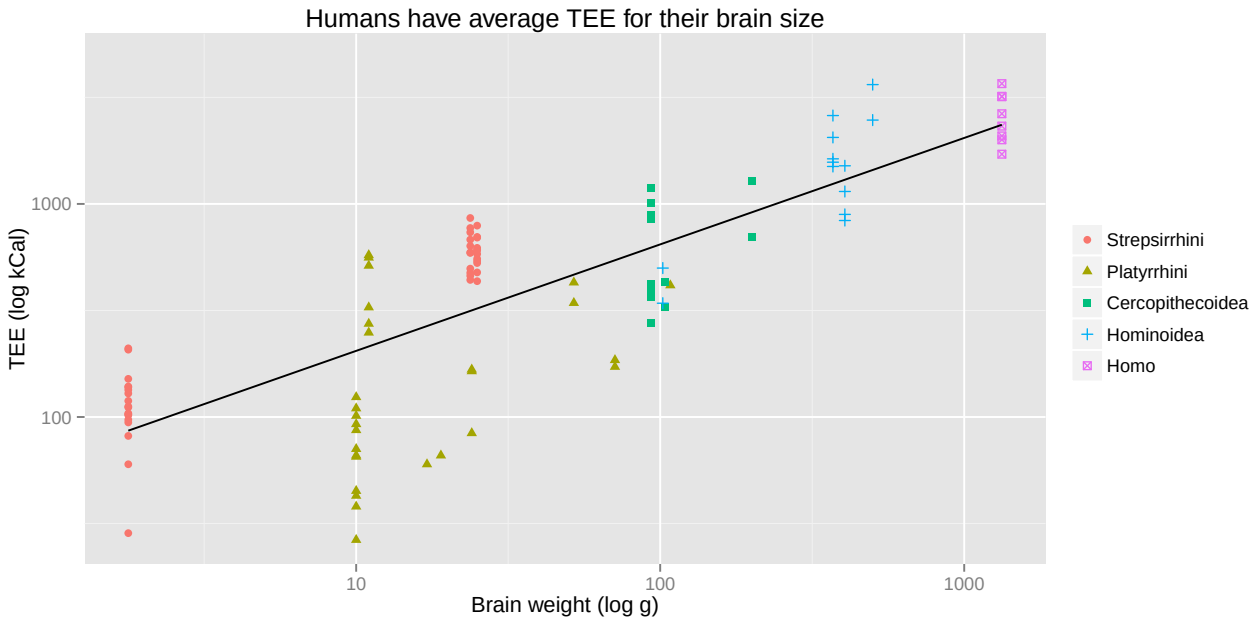


Model	Range	RMSE	MAE
W only	-0.6191508, 0.497406	0.2355909	0.1678582
Best model	-0.6040659, 0.443363	0.2256559	0.1490703
Key & Ross	-0.2077282, 1.2528322	0.4461273	0.133878
Leonard & Robertson	-0.2591463, 1.3386431	0.4751238	0.1222741

Humans do not have higher TEE than predicted by brain weight

a

```
## Warning: Removed 48 rows containing missing values (stat_smooth).

## Warning: Removed 93 rows containing missing values (geom_point).
```




```
## Warning: Removed 48 rows containing missing values (stat_smooth).
```

```
## Warning: Removed 93 rows containing missing values (geom_point).
```

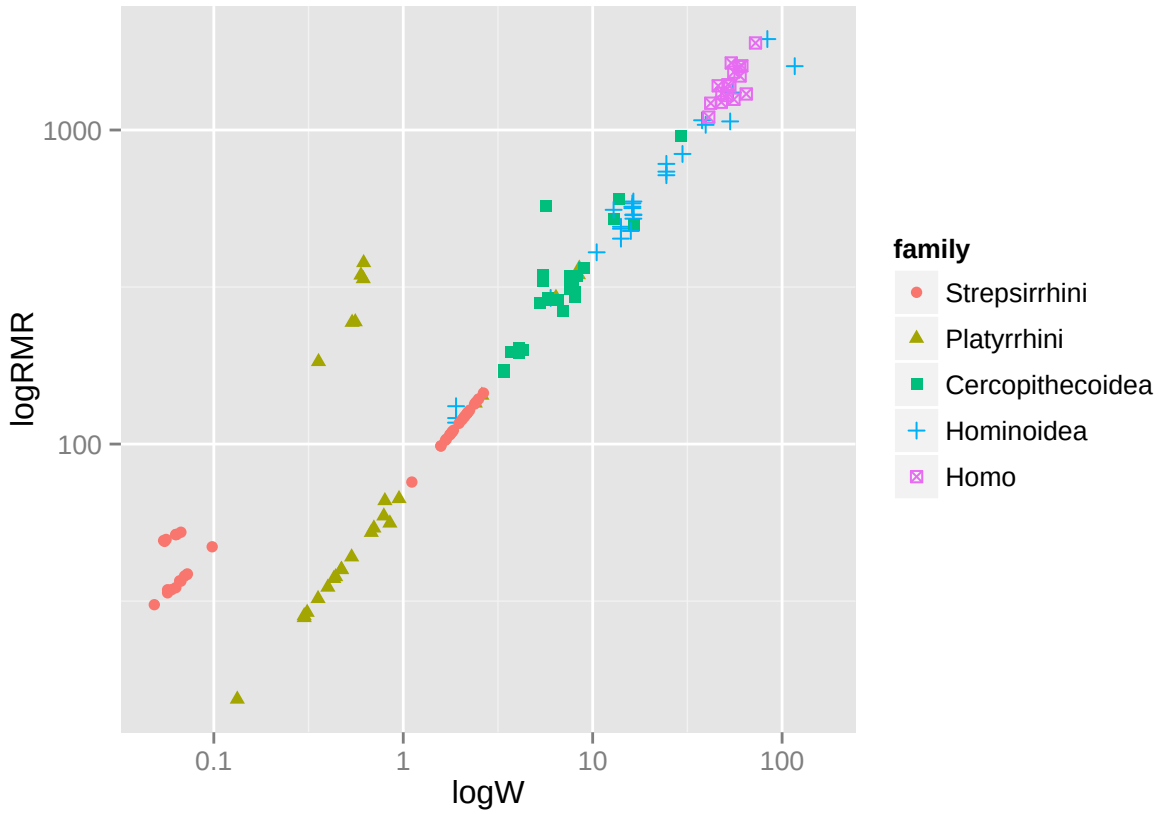
Supplementary Materials

RMR is not typically measured

Much of the data about RMR has been estimated from body weight, and not measured. This is readily apparent when looking at how closely the points fall to a straight line when plotted.

```
ggplot(data = data, aes(x = logW, y = logRMR)) +
  geom_point(data = data, aes(colour = family, shape=family)) +
  scale_x_continuous(breaks = c(-1,0,1,2,3,4), labels=c(0.1,1,10, 100,1000,10000)) +
  scale_y_continuous(breaks = c(-1,0,1,2,3,4), labels=c(0.1,1,10,100,1000,10000))
```

```
## Warning: Removed 52 rows containing missing values (geom_point).
```



This means that measurements are outliers

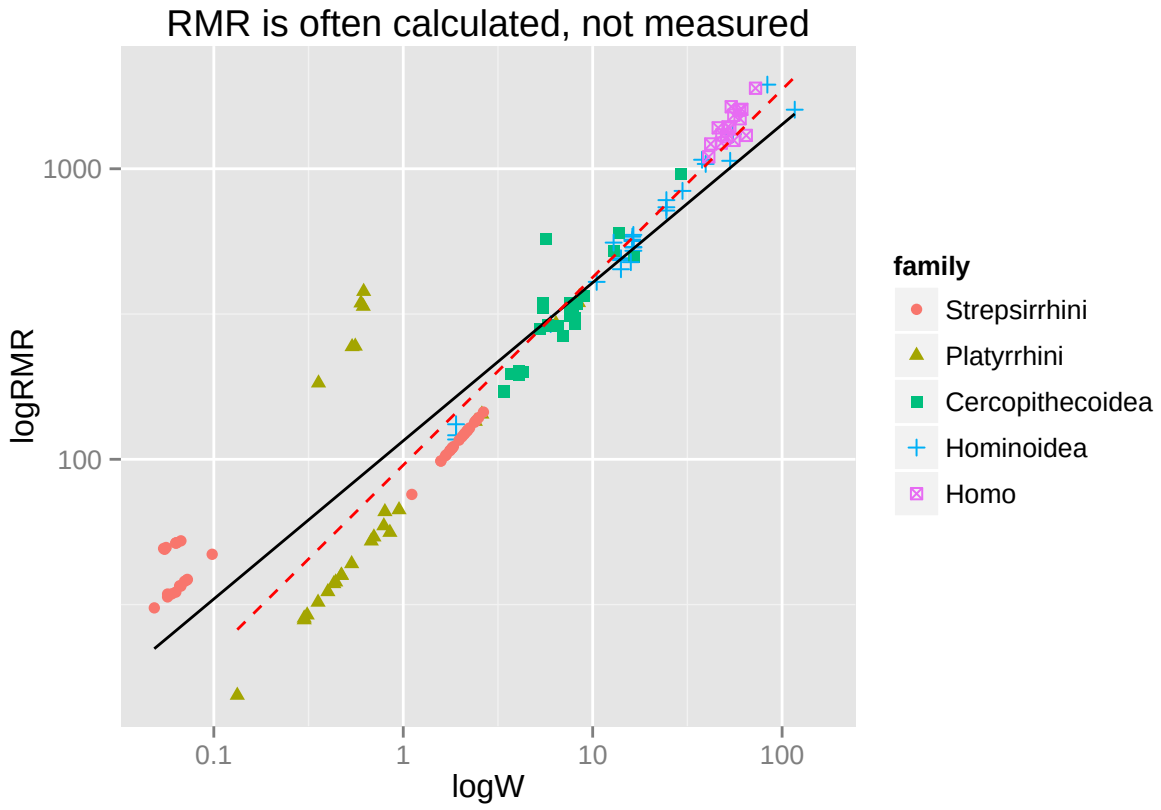
Including the actual measured data, which is often much higher, biases the trend line to a higher intercept and lower slope, and makes the human data appear ore divergent than it actually is. Removing the strepsirrhines, which have life history strategies very different from other primates, corrects for most of this problem.

```
ggplot(data = data, aes(x = logW, y = logRMR)) +
  geom_point(data = data, aes(colour = family, shape=family)) +
  stat_smooth(method = 'lm', colour = 'black', se=FALSE) +
  stat_smooth(data = data[!(data$family %in% c('Strepsirrhini')), ],
              method = 'lm', colour = 'red', se=FALSE, linetype='dashed')  +
  scale_x_continuous(breaks = c(-1,0,1,2,3,4), labels=c(0.1,1,10, 100,1000,10000)) +
  scale_y_continuous(breaks = c(-1,0,1,2,3,4), labels=c(0.1,1,10,100,1000,10000)) +
  ggtitle("RMR is often calculated, not measured")
```

```
## Warning: Removed 52 rows containing missing values (stat_smooth).
```

```
## Warning: Removed 50 rows containing missing values (stat_smooth).
```

```
## Warning: Removed 52 rows containing missing values (geom_point).
```




```
ggsave('suppRMR.png', width=18, height=14, units='cm')
```

```
## Warning: Removed 52 rows containing missing values (stat_smooth).
```

```
## Warning: Removed 50 rows containing missing values (stat_smooth).
```

```
## Warning: Removed 52 rows containing missing values (geom_point).
```

Data source has a large influence on TEE

```
##
## Call:
## lm(formula = logTEE ~ logW + Source, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.53513 -0.06251  0.00127  0.05872  0.29404
##
## Coefficients:
##                                                    Estimate
## (Intercept)                                         2.14973
## logW                                                0.75629
## SourceLeonard & Robertson 1997                     -0.15264
## SourceLeonard Galloway Ivakine 1997                -0.14520
## SourceLeonard Katzmarzyk Stephen Ross 1995         -0.05166
## SourcePontzer et al 2010                           -0.30467
```

```
## SourcePontzer et al 2014                                        -0.13000
## SourcePower Tardiff Power Layne 2003                             0.64814
## SourceRothman Dierenfeld Hintz Pell 2008                        -0.26042
## SourceRothman Dierenfeld Hintz Pell 2008, Smith Jungers 1997 -0.26164
## SourceSchmid Speakman 2000                                       0.79290
## SourceSimmen et al 2010                                          0.40097
## SourceWesterterp & Speakman 2008                                -0.10810
##                                                                 Std. Error
## (Intercept)                                                        0.02596
## logW                                                               0.02040
## SourceLeonard & Robertson 1997                                     0.03734
## SourceLeonard Galloway Ivakine 1997                                0.07154
## SourceLeonard Katzmarzyk Stephen Ross 1995                         0.07120
## SourcePontzer et al 2010                                           0.07971
## SourcePontzer et al 2014                                           0.04364
## SourcePower Tardiff Power Layne 2003                               0.05518
## SourceRothman Dierenfeld Hintz Pell 2008                           0.12544
## SourceRothman Dierenfeld Hintz Pell 2008, Smith Jungers 1997       0.12682
## SourceSchmid Speakman 2000                                         0.04739
## SourceSimmen et al 2010                                            0.03545
## SourceWesterterp & Speakman 2008                                   0.04623
##                                                                 t value
## (Intercept)                                                      82.825
## logW                                                             37.071
## SourceLeonard & Robertson 1997                                   -4.087
## SourceLeonard Galloway Ivakine 1997                              -2.030
## SourceLeonard Katzmarzyk Stephen Ross 1995                       -0.726
## SourcePontzer et al 2010                                         -3.822
## SourcePontzer et al 2014                                         -2.979
## SourcePower Tardiff Power Layne 2003                             11.747
## SourceRothman Dierenfeld Hintz Pell 2008                         -2.076
## SourceRothman Dierenfeld Hintz Pell 2008, Smith Jungers 1997     -2.063
## SourceSchmid Speakman 2000                                       16.732
## SourceSimmen et al 2010                                          11.311
## SourceWesterterp & Speakman 2008                                 -2.339
##                                                                 Pr(>|t|)
## (Intercept)                                                      < 2e-16 ***
## logW                                                             < 2e-16 ***
## SourceLeonard & Robertson 1997                                  7.15e-05 ***
## SourceLeonard Galloway Ivakine 1997                             0.044208 *
## SourceLeonard Katzmarzyk Stephen Ross 1995                      0.469258
## SourcePontzer et al 2010                                        0.000195 ***
## SourcePontzer et al 2014                                        0.003382 **
## SourcePower Tardiff Power Layne 2003                             < 2e-16 ***
## SourceRothman Dierenfeld Hintz Pell 2008                        0.039621 *
## SourceRothman Dierenfeld Hintz Pell 2008, Smith Jungers 1997 0.040866 *
## SourceSchmid Speakman 2000                                       < 2e-16 ***
## SourceSimmen et al 2010                                          < 2e-16 ***
## SourceWesterterp & Speakman 2008                                0.020707 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1173 on 147 degrees of freedom
##   (35 observations deleted due to missingness)
```

```
## Multiple R-squared:  0.963,  Adjusted R-squared:   0.96
## F-statistic: 318.7 on 12 and 147 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = logTEE ~ logW + CAN, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.67039 -0.09548  0.03094  0.13821  0.42591
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.48304    0.02333  106.42  < 2e-16 ***
## logW         0.48867    0.01675   29.17  < 2e-16 ***
## CAN         -0.27710    0.05238   -5.29 4.04e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2191 on 157 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.8621, Adjusted R-squared:  0.8603
## F-statistic: 490.7 on 2 and 157 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = logTEE ~ logW + LR, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5922 -0.1201  0.0126  0.1666  0.4741
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.46739    0.02258 109.267  < 2e-16 ***
## logW         0.51580    0.01664  30.989  < 2e-16 ***
## LR          -0.22799    0.04561  -4.998 1.53e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2209 on 157 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.8598, Adjusted R-squared:  0.858
## F-statistic: 481.4 on 2 and 157 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = logTEE ~ logW + P10, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.61963 -0.12369  0.03152  0.17283  0.49966
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.43371    0.02310 105.340   <2e-16 ***
## logW         0.50901    0.01794  28.372   <2e-16 ***
## P10         -0.13258    0.13964  -0.949    0.344
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2372 on 157 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.8384, Adjusted R-squared:  0.8364
## F-statistic: 407.3 on 2 and 157 DF,  p-value: < 2.2e-16

##
## Call:
## lm(formula = logTEE ~ logW + P14, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.62991 -0.12380  0.03897  0.17013  0.49168
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.44413    0.02320 105.372   <2e-16 ***
## logW         0.51104    0.01758  29.064   <2e-16 ***
## P14         -0.15406    0.06368  -2.419   0.0167 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2335 on 157 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.8433, Adjusted R-squared:  0.8413
## F-statistic: 422.6 on 2 and 157 DF,  p-value: < 2.2e-16

##
## Call:
## lm(formula = logTEE ~ logW + POW, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.61915 -0.12251  0.03679  0.17400  0.49741
##
## Coefficients: (1 not defined because of singularities)
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.43306    0.02309  105.39   <2e-16 ***
## logW         0.50659    0.01775   28.54   <2e-16 ***
## POW               NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2371 on 158 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.8375, Adjusted R-squared:  0.8365
## F-statistic: 814.3 on 1 and 158 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = logTEE ~ logW + SIM, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.57453 -0.12213  0.02595  0.11387  0.55768
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.38941    0.02387 100.121  < 2e-16 ***
## logW         0.52044    0.01704  30.534  < 2e-16 ***
## SIM          0.23051    0.05129   4.494 1.35e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2239 on 157 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.856,  Adjusted R-squared:  0.8542
## F-statistic: 466.7 on 2 and 157 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = logTEE ~ logW + WS, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.61915 -0.12251  0.03679  0.17400  0.49741
##
## Coefficients: (1 not defined because of singularities)
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.43306    0.02309  105.39   <2e-16 ***
## logW         0.50659    0.01775   28.54   <2e-16 ***
## WS                NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2371 on 158 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.8375, Adjusted R-squared:  0.8365
## F-statistic: 814.3 on 1 and 158 DF,  p-value: < 2.2e-16
```

But this is expected based on the species in each study

```
## Using WS as value column: use value.var to override.
## Aggregation function missing: defaulting to length
```

```
##                                                Source Strepsirrhini
## 1                                           CAN 2003             0
## 2                                         Coehlo 1986            0
## 3                                Leonard & Robertson 1997          0
## 4                           Leonard Galloway Ivakine 1997          0
## 5                     Leonard Katzmarzyk Stephen Ross 1995         0
## 6                                   Pontzer et al 2010             0
## 7                                   Pontzer et al 2014             2
## 8                          Power Tardiff Power Layne 2003          0
## 9                        Rothman Dierenfeld Hintz Pell 2008        0
## 10 Rothman Dierenfeld Hintz Pell 2008, Smith Jungers 1997          0
## 11                                    Schmid Speakman 2000        18
## 12                                      Simmen et al 2010         23
## 13                               Westerterp & Speakman 2008        0
##    Platyrrhini Cercopithecoidea Hominoidea Homo
## 1           12                9          2    0
## 2            0               17         16    0
## 3           10                5          6    8
## 4            0                0          0    4
## 5            0                0          0    4
## 6            0                0          3    0
## 7            2                5          4    2
## 8            6                0          0    0
## 9            0                0          1    0
## 10           0                0          1    0
## 11           0                0          0    0
## 12           0                0          0    0
## 13           0                0          0   35
```

```
## Warning in chisq.test(data$family, data$Source): Chi-squared approximation
## may be incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  data$family and data$Source
## X-squared = 433.55, df = 48, p-value < 2.2e-16
```

```
## Using WS as value column: use value.var to override.
## Aggregation function missing: defaulting to length
```

Factorial method underestimates energy expenditure

But this may be a result of its interaction with study population - see preceeding section

```
##
## Call:
## lm(formula = logTEE ~ logW + Method, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.62763 -0.10656 -0.00143  0.13656  0.42514
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 2.51314    0.02264 110.987  < 2e-16 ***
## logW                        0.51310    0.01487  34.498  < 2e-16 ***
## MethodFactorial            -0.28166    0.03292  -8.557 1.06e-14 ***
## MethodHeart Rate           -0.03420    0.07170  -0.477    0.634
## MethodRespiratory Quotient  0.21931    0.08163   2.687    0.008 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1904 on 155 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.8971, Adjusted R-squared:  0.8945
## F-statistic: 337.9 on 4 and 155 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = logTEE ~ logW + FA, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.64761 -0.09851 -0.00707  0.13893  0.43049
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.52993    0.02177 116.207  < 2e-16 ***
## logW         0.50389    0.01451  34.721  < 2e-16 ***
## FA          -0.29174    0.03271  -8.919 1.15e-15 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1938 on 157 degrees of freedom
##   (35 observations deleted due to missingness)
## Multiple R-squared:  0.8921, Adjusted R-squared:  0.8908
## F-statistic: 649.3 on 2 and 157 DF,  p-value: < 2.2e-16
```

Human TEE effect is robust to removing all factorial data

```
## Start:  AIC=-388.79
## logTEE ~ logW
##
##         Df Sum of Sq    RSS     AIC
## + human  1   0.13557 2.7085 -392.06
## <none>               2.8440 -388.79
##
## Step:  AIC=-392.06
## logTEE ~ logW + human

##
## Call:
## lm(formula = logTEE ~ logW + human, data = data[data$Method !=
##     "Factorial", ])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.70328 -0.05873  0.00202  0.08684  0.29194
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.55400    0.02025 126.144   <2e-16 ***
## logW         0.41278    0.02093  19.722   <2e-16 ***
```

```
## humanHuman   0.11158    0.04867   2.293   0.0239 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1606 on 105 degrees of freedom
##   (33 observations deleted due to missingness)
## Multiple R-squared:  0.9141, Adjusted R-squared:  0.9124
## F-statistic: 558.5 on 2 and 105 DF,  p-value: < 2.2e-16
```

Human TEE effect is robust to removing the respiratory quotient data

```
## Start:  AIC=-447.24
## logTEE ~ logW
##
##         Df Sum of Sq    RSS     AIC
## + human  1   0.74938 7.4725 -459.96
## <none>               8.2219 -447.24
##
## Step:  AIC=-459.96
## logTEE ~ logW + human

##
## Call:
## lm(formula = logTEE ~ logW + human, data = data[data$Method !=
##     "Respiratory Quotient", ])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5826 -0.1295 -0.0003  0.1679  0.4764
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.39294    0.02305 103.806  < 2e-16 ***
## logW         0.45573    0.02324  19.610  < 2e-16 ***
## humanHuman   0.20099    0.05165   3.891 0.000149 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2225 on 151 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.8621, Adjusted R-squared:  0.8603
## F-statistic: 472.1 on 2 and 151 DF,  p-value: < 2.2e-16
```

Sampling is not independent

Because:

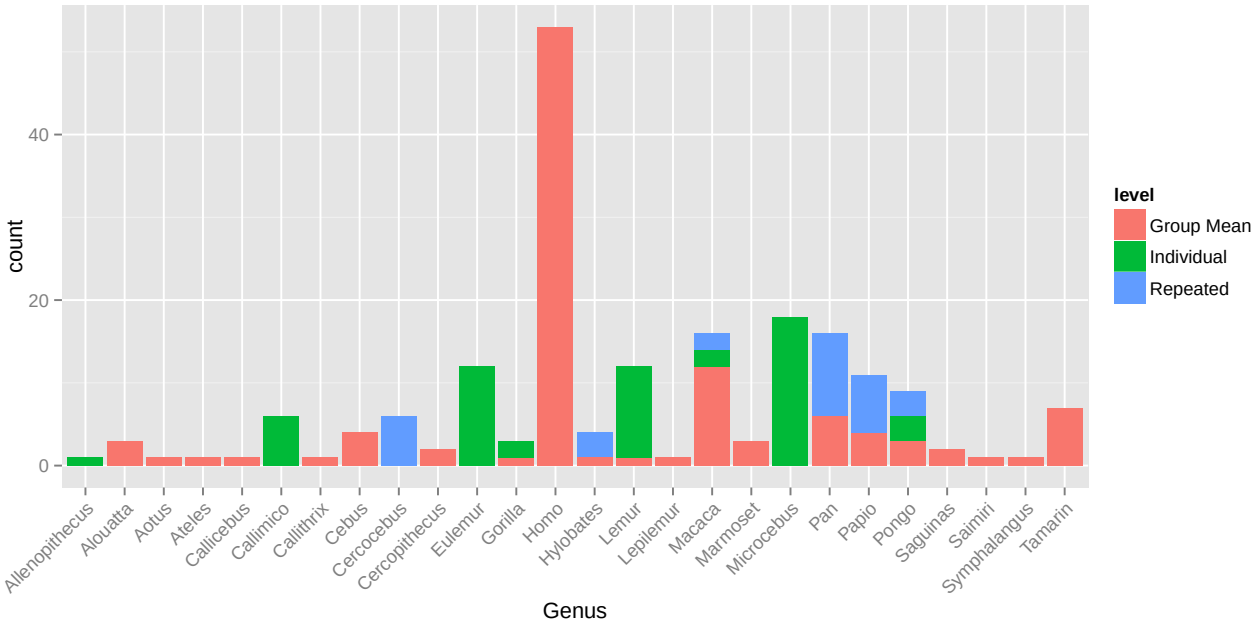
1. Species are unevenly represented; and,
2. Species will tend give similar results to close phylogenetic relations
3. Data exist at three levels of measurement

Explicitly including this in the model is difficult because:

1. Controlling for phylogeny will destroy the phylogenetic differences that are the property under investigation; and

2. Many genera are only attested at one hierarchy

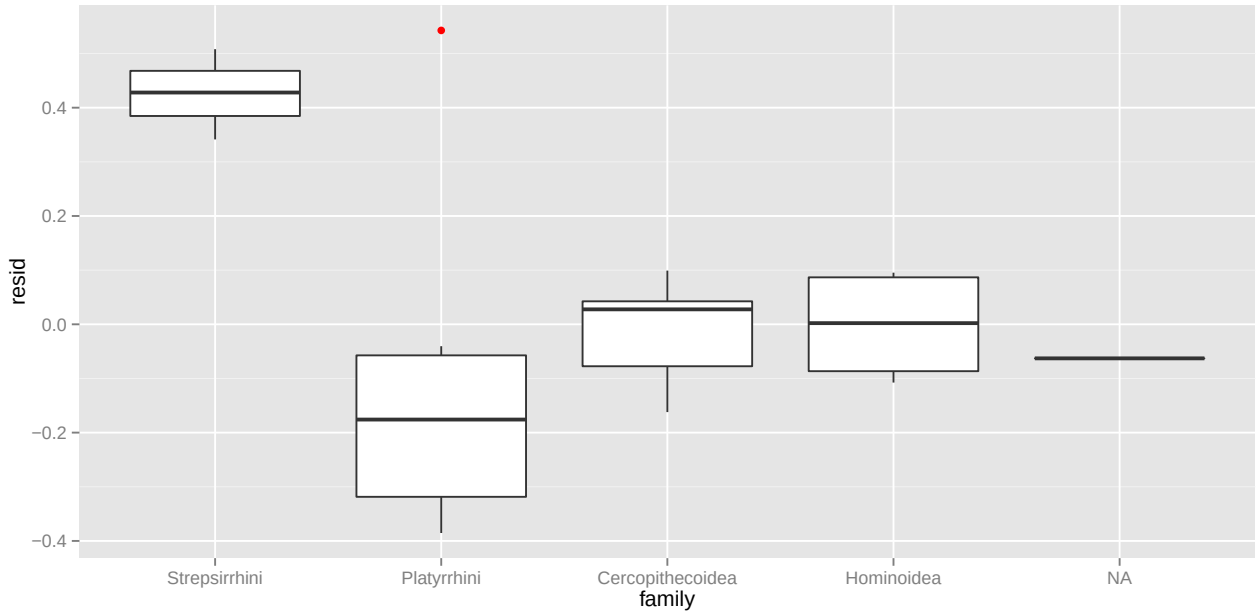
```
##
##  Chi-squared test for given probabilities
##
## data:  table(data$Genus)
## X-squared = 382.47, df = 25, p-value < 2.2e-16
```



Aggregate results show the same pattern for genera differences

```
data.species <- merge(aggregate(logW ~ Species,
                                data = data[is.na(data$logTEE) == FALSE,],
                                FUN = mean),
                      aggregate(logTEE ~ Species,
                                data = data[is.na(data$logTEE) == FALSE,],
                                FUN = mean))
data.species <- na.omit(data.species)
data.species$human <- factor(ifelse(data.species$Species == 'sapiens', 1, 0),
                             labels = c("Nonhuman primate","Human"))
data.species <- merge(aggregate(logRMR ~ Species,
                                data = data[is.na(data$logRMR) == FALSE,],
                                FUN = mean),
                      data.species)
data.species <- merge(aggregate(PAL ~ Species,
                                data = data[is.na(data$PAL) == FALSE,],
                                FUN = mean),
                      data.species)
```

```
data.species$family <- factor(ifelse(data.species$human == "Human", 4, NA),
                              labels = c("Strepsirrhini","Platyrrhini","Cercopithecoidea","Hominoidea",
                              levels = c(0,1,2,3,4))
data.species$family[data.species$Species %in% c("lar", "troglodytes", "pygmaeus", "syndactylus", "bering
data.species$family[data.species$Species %in% c("murinus", "catta", "sp", "ruficaudatus")] <- "Strepsirr
data.species$family[data.species$Species %in% c("palliata", "trivirgatus", "geoffroyi", "moloch", "apell
data.species$family[data.species$Species %in% c("Albigena", "guereza", "fascicularis", "anubis", "mulatt
data.species$gorilla <- factor(ifelse(data.species$Species %in% c('beringei', 'gorilla') ,1,0))
data.species$pan <- factor(ifelse(data.species$Species %in% c("troglodytes", "paniscus") ,1,0))
data.species$pongo <- factor(ifelse(data.species$Species == "pygmaeus",1,0))
data.species$hylobates <- factor(ifelse(data.species$Species %in% c("syndactylus", "lar"),1,0))
```




```
##
## Call:
## lm(formula = logTEE ~ logW + family, data = data.species)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.24653 -0.09542 -0.01144  0.07038  0.68756
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.27399    0.09487  23.969 1.52e-14 ***
## logW         0.59378    0.09595   6.189 9.91e-06 ***
## family1     -0.36127    0.14467  -2.497   0.0231 *
## family2      0.11708    0.11951   0.980   0.3410
## family3      0.06250    0.10605   0.589   0.5634
## family4      0.09226    0.11835   0.780   0.4463
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2074 on 17 degrees of freedom
##   (1 observation deleted due to missingness)
```

```
## Multiple R-squared:  0.8688, Adjusted R-squared:  0.8302
## F-statistic: 22.52 on 5 and 17 DF,  p-value: 6.031e-07
```

Aggregate results show the same cause of variability in apes

```
summary(lm(logTEE ~ logW + pan, data = data.species))
```

```
##
## Call:
## lm(formula = logTEE ~ logW + pan, data = data.species)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.38625 -0.15748 -0.05141  0.10153  0.53996
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.20351    0.06662  33.078  < 2e-16 ***
## logW         0.55406    0.07771   7.130 4.96e-07 ***
## pan1         0.10994    0.28615   0.384    0.705
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2665 on 21 degrees of freedom
## Multiple R-squared:  0.7349, Adjusted R-squared:  0.7096
## F-statistic: 29.11 on 2 and 21 DF,  p-value: 8.83e-07
```

```
#summary(lm(logTEE ~ logW + gorilla, data = data.species))
summary(lm(logTEE ~ logW + pongo, data = data.species))
```

```
##
## Call:
## lm(formula = logTEE ~ logW + pongo, data = data.species)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.38701 -0.15774 -0.05207  0.10672  0.53908
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.20423    0.06671  33.040  < 2e-16 ***
## logW         0.55345    0.07929   6.980 6.82e-07 ***
## pongo1       0.10049    0.29196   0.344    0.734
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2666 on 21 degrees of freedom
## Multiple R-squared:  0.7345, Adjusted R-squared:  0.7092
## F-statistic: 29.05 on 2 and 21 DF,  p-value: 8.959e-07
```

```r
summary(lm(logTEE ~ logW + hylobates, data = data.species))
```

```
##
## Call:
## lm(formula = logTEE ~ logW + hylobates, data = data.species)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.39050 -0.17411 -0.02973  0.08224  0.53875
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.20884    0.06724  32.852  < 2e-16 ***
## logW         0.56938    0.07460   7.632 1.74e-07 ***
## hylobates1  -0.10460    0.19860  -0.527    0.604
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2656 on 21 degrees of freedom
## Multiple R-squared:  0.7365, Adjusted R-squared:  0.7114
## F-statistic: 29.35 on 2 and 21 DF,  p-value: 8.281e-07
```

```r
summary(lm(logTEE ~ logW + human, data = data.species))
```

```
##
## Call:
## lm(formula = logTEE ~ logW + human, data = data.species)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.38925 -0.14913 -0.05152  0.11652  0.53417
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.20552    0.06585  33.493  < 2e-16 ***
## logW         0.54001    0.07830   6.897 8.15e-07 ***
## humanHuman   0.23715    0.28830   0.823     0.42
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2632 on 21 degrees of freedom
## Multiple R-squared:  0.7414, Adjusted R-squared:  0.7167
## F-statistic:  30.1 on 2 and 21 DF,  p-value: 6.812e-07
```

Aggregate results show the same human TEE effect

The coefficient is no longer statistically significant, but we believe that this is only a function of the reduction in sample size, as the coefficient itself has actually gotten larger

```r
summary(lm(logTEE ~ logW + human, data=data.species))
```

```
##
## Call:
## lm(formula = logTEE ~ logW + human, data = data.species)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.38925 -0.14913 -0.05152  0.11652  0.53417
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.20552    0.06585  33.493  < 2e-16 ***
## logW         0.54001    0.07830   6.897 8.15e-07 ***
## humanHuman   0.23715    0.28830   0.823     0.42
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2632 on 21 degrees of freedom
## Multiple R-squared:  0.7414, Adjusted R-squared:  0.7167
## F-statistic:  30.1 on 2 and 21 DF,  p-value: 6.812e-07
```

```
print.xtable(xtable(summary(lm(logTEE ~ logW + human, data=data.species))), file='supptee.tex')
ggplot(data = data.species, aes(x = logW, y = logTEE)) +
  geom_point(aes(colour = family, shape = family), size=3) +
  stat_smooth(data = data[data$human != 'Human',], method = 'lm',
              colour = 'black', se=FALSE, fullrange=TRUE) +
  scale_x_continuous(breaks = c(-1,0,1,2,3,4), labels=c(0.1,1,10, 100,1000,10000)) +
  scale_y_continuous(breaks = c(-1,0,1,2,3,4), labels=c(0.1,1,10,100,1000,10000)) +
  ggtitle("Humans are farther from trend line
          in aggregated data")
```

```
## Warning: Removed 35 rows containing missing values (stat_smooth).
```

```
## Warning: Removed 1 rows containing missing values (geom_point).
```


```
ggsave('suppape.png', width=18, height=14, units='cm')
```

```
## Warning: Removed 35 rows containing missing values (stat_smooth).
```

```
## Warning: Removed 1 rows containing missing values (geom_point).
```