Heart rate variability in mental stress: The data reveal regression to the mean

Prateek
Somendra Singh Kharola
Divyanshu Gupta
Anshuman Agrawal

Indian Statistical Instiute
Bangalore centre

April 2023

ABSTRACT

The aim of the original work, by D.A. Dimitriev et al. titled "Heart rate variability in mental stress: The data reveal regression to the mean", is to assess whether there is any relationship between heart rate variability(HRV) and mental stress. Before and while applying mental stress to participants, their HRV was recorded. It was observed that regression to the mean is a major source of variability of stress-related changes in heart rate variability.

Contents

1. Introduction

Before we can explain the research study, there are a few terms to know about.

1.1 Regression to the mean

Regression toward the mean(RTM) (also called reversion to the mean, and reversion to mediocrity) is the phenomenon where if one sample of a random variable is extreme, the next sampling of the same random variable is likely to be closer to its mean[1].

The phenomenon of Regression Toward the Mean (RTM) can create the appearance of genuine change in repeated data, even when the variation is actually due to natural fluctuations. This occurs because extreme measurements, whether they are unusually high or low, tend to be followed by measurements that are closer to the overall mean.

It happens because values are observed with random error. Random errors are non-systematic variation in the observed values around a true mean. It is uncommon to obtain data without random error, which leads to the prevalence of the Regression Toward the Mean (RTM) phenomenon.

1.2 Heart rate variability

Heart rate variability is the physiological phenomenon of variation in the time interval between heartbeats[2]. "A healthy heart is not a metronome"[3]. Heart beats are complex and non-linear. A healthy heart's beat-to-beat fluctuations are best described by mathematical chaos. Therefore, a number of measures have been developed to quantify and better understand this 'chaos'.

1.2.1 HRV metrics

There are a number of ways to measure HRV. Some of them are:-

Parameter	Unit
SDNN	ms
SDRR	ms
SDANN	ms
SDNNI	ms
pNN50	%
TINN	ms
RMSSD	ms

Table 1.1: Time domain measurements

Parameter	Unit
ULF power	ms^2
VLF power	ms^2
LF peak	Hz
LF power	ms^2
HF peak	Hz
HF power	ms^2
LF/HF	%

Table 1.2: Frequency domain measurements

Parameter	Unit
S	ms
SD1	ms
SD2	ms
SD1/SD2	%
ApEn	-
SampEn	-
D_2	-

Table 1.3: Non-linear measurements

The research paper in question focuses on five measures, namely:

1. Heart rate(in bpm):-
   Heart rate (or pulse rate) is the frequency of the heartbeat measured by the number of contractions of the heart per minute (beats per minute, or bpm).[4]

2. SDNN(in ms):-
   The standard devsation of the interbeat interval of normal sinus beats (SDNN) is measured in ms. "Normal" means that abnormal beats, like ectopic beats (heartbeats that originate outside the right atrium's sinoatrial node), have been removed.

3. LF power(in ms^2):-
   Power is the signal energy found within a frequency band. Low frequency is considered to be of the range(0.04-0.15 Hz).This region was previously called the baroreceptor range because it mainly reflects baroreceptor activity during resting conditions.

4. HF power(in ms^2):-
   The HF or respiratory band (0.15-0.40 Hz) is conventionally recorded over a minimum 1 min period. The HF band reflects parasympathetic activity and is called the respiratory band because it corresponds to the HR variations related to the respiratory cycle.

5. LF/HF(in %):-
   It is just the ratio of LF and HF.

The research study in question suggests that when an individual with high/low baseline HRV is subjected to stress, the change in HRV is not entirely due to stress, but majorly due to regression to the mean. Although it may appear that mental stress is causing HRV, it is actually regression to the mean that is doing so.

2. Data collection

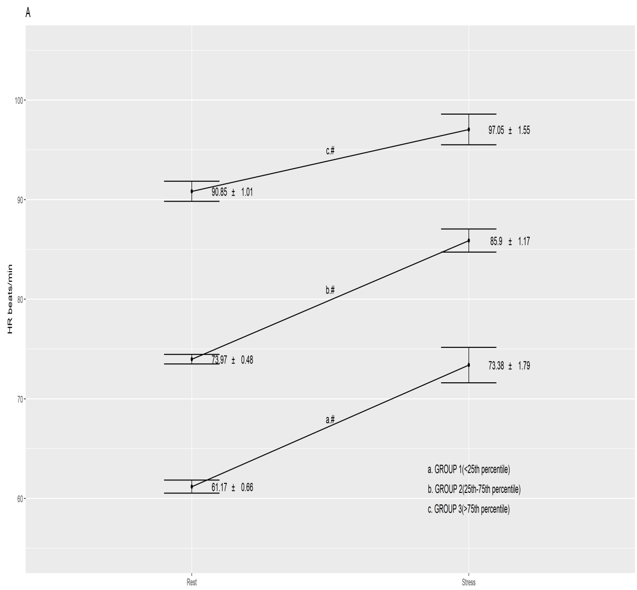
A total of 1156 students attending Chuvash State Pedagogical University, Russia were considered for participation in this study of heart rate variability. Out of 1156, 162 students were randomly selected and their baseline ECG recordings were acquired. Using Kubios heart rate variability software with the raw ECG recording as input, the HRV parameters of the participants were acquired. The parameters can then be saved as a pdf or matlab file.

Since the data did not follow normal distribution, the HRV measurements had to be log transformed to obtain normal distribution.

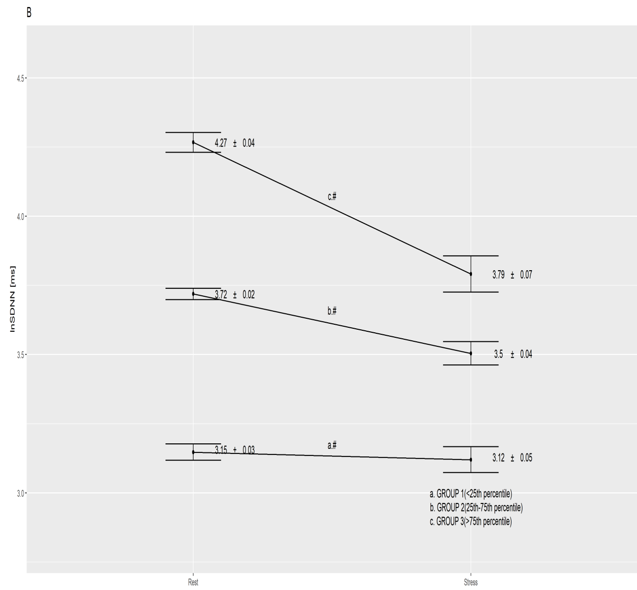
The participants were then divided into three groups based on their baseline measurements. One group with baseline measurements in first quartile. Another with baseline measurements in forth quartile and the rest in third group. This was done for all five measurements.

Then, the participants were subjected to a mental stress stimulus, which in this case was 10 minutes of serial subtraction of 7 from random 3 digit numbers. The HRV was again measured during the stress-test.

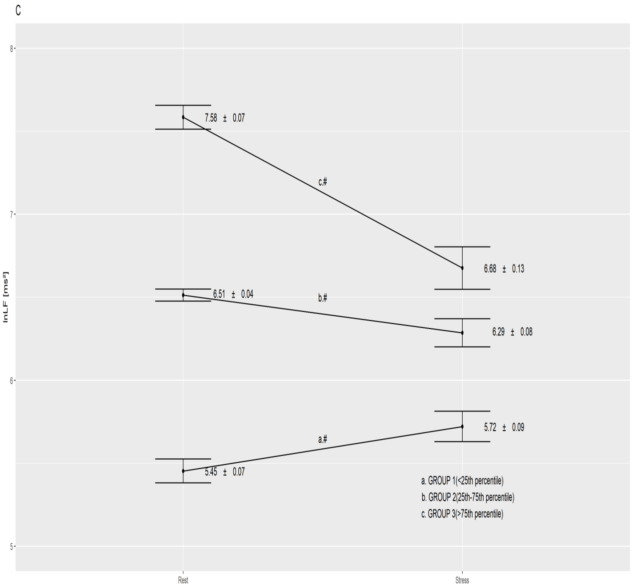
The measurements were then plotted against the two categories(rest and stress) and the results are as follows:



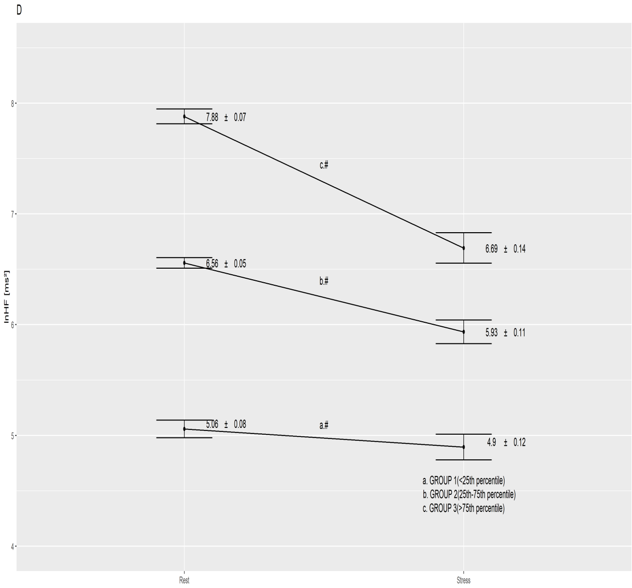
(a) Heart rate rest vs stress



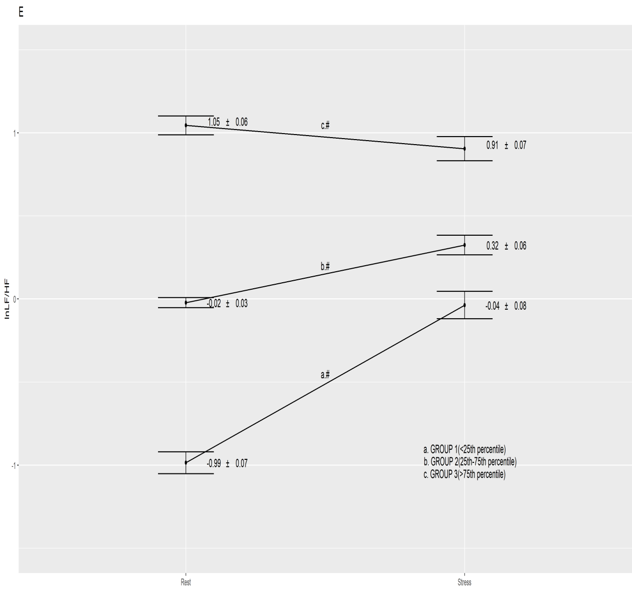
(b) LnSDNN rest vs stress



(c) LnLF rest vs stress



(d) LnHF rest vs stress



(e) LnLF/HF rest vs stress

3. Observation

In each figure, mean of each group's measurements at rest is joined with mean of each group's measurements during stress. The bar signifies one standard error.
As one would expect, the heart rate increases when the participants were subjected to stress. However, the change in the heart rate is significant in group a and b but insignificant in group c.

The LnSDNN changes were significant in group c and b but not in the group with low baseline level.

The passage from rest to stress evoked a significant decrease in group c, but group a and b exhibited insignificant changes.

Statistical analysis of LnHF revealed a significant decrease in this parameter in the groups b and c and insignificant changes of LnHF were observed in the group a.

During stress, the LnLF/HF increased significantly in group a while group b and c showed insignificant changes.

In order to verify whether these groups are behaving differently, we used ANOVA. However, before using ANOVA, its assumptions need to be verified.
ANOVA uses three assumptions:

1. Normality:- the data is normally distributed.
2. Independence:- The observations are independent of each other.
3. Homogeneity of variance:- The difference between variances of different groups should be insignificant.

The data was log-transformed to be normally distributed. As the experiment's design suggests, the observations are independent of each other since one person can't affect other person's HRV. Levene's test was used to verify homogeneity of variance.

3.1 Levene's test

Levene's test works as follows:
Null hypothesis:- The variances of all groups are equal.
Alternative hypothesis:- The variances of at least two groups are different.
Test statistic:- Given a variable Y with sample of size N divided into k subgroups, where N_i is the sample size of the ith subgroup, the Levene test statistic is defined as:

$$W = \frac{N-k}{k-1} \frac{\sum_{i=1}^{k} N_i (Z_{i\cdot} - Z_{\cdot\cdot})^2}{\sum_{i=1}^{k} \sum_{j=1}^{N_i} (Z_{ij} - Z_{i\cdot})}$$

where,

$$Z_{ij} = |Y_{ij} - \bar{Y}_i|, \bar{Y}_i \text{ is a mean of } i\text{-th group}$$
$$Z_{i\cdot} = \frac{1}{N_i}\sum_{j=1}^{N_i} Z_{ij}$$
$$Z_{\cdot\cdot} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{N_i} Z_{ij}$$

W approximately follows F-distribution with degrees of freedom $k-1, N-k$.
We reject the null hypothesis if $W > F_{k-1,N-k}(1-\alpha)$ where F is a quantile function of F-distribution with significance level α and degrees of freedom $k-1, N-k$.

Applying levene's test in our data at 5% significance level, we found out that the difference of variance is indeed insignificant in all five parameters. Thus, we can use ANOVA without any problems.

3.2 ANOVA results

Applying ANOVA for all five measurements, we get the following results:

1. ANOVA for heart rate:-
   Null hypothesis:- Mean of change of heart rate is same in all three groups.
   Alternative hypothesis:- Mean of change of heart rate of at least two groups differs.

	Sum of squares	df	mean of squares	F-ratio	p-value
Between groups	1060.53	2	530.27	5.46	0.005
Within groups(Error)	15452	159	97.2		
Total	16512.38	161			

   Since p -value$<$ 0.05, we reject the null hypothesis.

2. ANOVA for LnSDNN:-
   Null hypothesis:- Mean of change of lnSDNN is same in all three groups.
   Alternative hypothesis:- Mean of change of lnSDNN of at least two groups differs.

	Sum of squares	df	mean of squares	F-ratio	p-value
Between groups	4.13	2	2.06	15.6	6.3×10^{-7}
Within groups(Error)	21	159	0.132		
Total	25.13	161			

   Since p -value$<$ 0.05, we reject the null hypothesis.

3. ANOVA for LnLF-
   Null hypothesis:- Mean of change of lnLF is same in all three groups.
   Alternative hypothesis:- Mean of change of lnLF of at least two groups differs.

	Sum of squares	df	mean of squares	F-ratio	p-value
Between groups	28.40	2	14.20	25.4	2.77×10^{-10}
Within groups(Error)	89.1	159	0.56		
Total	117.50	161			

   Since p -value< 0.05, we reject the null hypothesis.

4. ANOVA for LnHF:-
   Null hypothesis:- Mean of change of lnHF is same in all three groups.
   Alternative hypothesis:- Mean of change of lnHF of at least two groups differs.

	Sum of squares	df	mean of squares	F-ratio	p-value
Between groups	21.37	2	10.69	13.2	4.95×10^{-6}
Within groups(Error)	129	159	0.81		
Total	150.37	161			

   Since p -value< 0.05, we reject the null hypothesis.

5. ANOVA for LnLF/HF:-
   Null hypothesis:- Mean of change of lnLF/HF is same in all three groups.
   Alternative hypothesis:- Mean of change of lnLF/HF of at least two groups differs.

	Sum of squares	df	mean of squares	F-ratio	p-value
Between groups	24.2	2	12.1	43.9	6.66×10^{-16}
Within groups(Error)	43.8	159	0.276		
Total	68	161			

   Since p -value< 0.05, we reject the null hypothesis.

3.3 Post hoc tests

Since ANOVA only gives information about equality of all group's measurements, we still don't know which group's measurements is different from the rest. In order to know this, we use post hoc tests. However, before we understand post hoc tests, we must go over some terminologies:

1. Statistical power:-
   Statistical power is the probability of rejecting a false null hypothesis. While performing a statistical test, we would like to reduce the probability of committing type I and type II errors. We can reduce the probability of committing a type I error by simply reducing our significance level but this increases the risk of committing a type II error.

2. FamilyWise Error Rate:-
   Family Wise Error Rate(FWER) is the probability of committing at least one type I error while doing multiple tests. Its value is given by

$$FWER = 1 - (1 - \alpha)^k$$

where

$$k = \text{number of comparisons}$$
$$\alpha = \text{significance level}$$

The value of FWER increases exponentially thus we can't use multiple t-tests to check which group's measurements differ in case ANOVA rejects the null hypothesis. This is why post hoc tests are needed.

3.3.1 Post hoc Bonferroni test

The Bonferroni test is a straightforward post hoc analysis method that involves conducting a series of t-tests on each pair of groups. As we previously mentioned, the number of groups can rapidly increase the number of comparisons, leading to inflated Type I error rates. To counteract this, the Bonferroni test divides the significance level α by the number of comparisons made, ensuring that the sum of the tests maintains the original Type I error rate.FWER for this test is given by

$$FWER = 1 - \left(1 - \frac{\alpha}{k}\right).$$

Another way of applying post hoc Bonferroni test is to keep the significance level same and multiplying p-values acquired using Fisher's LSD test or multiple t-tests by number of comparisons made and reducing p-values greater than one to 1.
The results from it are as follows:

1. Heart rate:-

	t-statistic	p-value
Group a vs group b	0.15	$2.61 \rightarrow 1$
Group b vs group c	3.01	0.009
group c vs group a	-2.75	0.02

Table 3.1: p-values received after applying post hoc Bonferroni on heart rate ANOVA.

Hence, change in heart rate when participants from group b are subjected to mental stress is significantly different compared to group c at 5% significance level. Similar can be said for group c and group a.

2. LnSDNN:-

	t-statistic	p-value
Group a vs group b	2.7	0.02
Group b vs group c	3.01	0.0008
group c vs group a	-2.75	3.44×10^{-7}

Table 3.2: p-values received after applying post hoc Bonferroni on lnSDNN ANOVA.

Hence, change in LnSDNN is significantly different in each group when participants are subjected to mental stress.

	t-statistic	p-value
Group a vs group b	3.46	0.002
Group b vs group c	4.71	1.61×10^{-5}
group c vs group a	-7.08	1.34×10^{-10}

Table 3.3: p-values received after applying post hoc Bonferroni on lnLF ANOVA.

3. LnLF:-

   Hence, change in LnLF is significantly different in each group when participants are subjected to mental stress.

4. LnHF:-

	t-statistic	p-value
Group a vs group b	2.67	0.025
Group b vs group c	3.25	0.0043
group c vs group a	-5.13	2.55×10^{-6}

Table 3.4: p-values received after applying post hoc Bonferroni on lnHF ANOVA.

   Hence, change in LnHF is significantly different in each group when participants are subjected to mental stress.

5. LnLF/HF:-

	t-statistic	p-value
Group a vs group b	5.98	4.35×10^{-8}
Group b vs group c	4.83	9.68×10^{-6}
group c vs group a	-9.39	3.44×0

Table 3.5: p-values received after applying post hoc Bonferroni on lnLF/HF ANOVA.

   Hence, change in LnLF/HF is significantly different in each group when participants are subjected to mental stress.

3.3.2 Post hoc Holm's test

The Bonferroni test has lower statistical power. This is due to several reasons, including the high probability of committing Type II error rates for each test. In other words, the Bonferroni test overcompensates for Type I errors, resulting in reduced power.

The Holm–Bonferroni method,also called the Holm method or Bonferroni–Holm method, is used to counteract the problem of multiple comparisons. It is intended to control the family-wise error rate (FWER) and offers a simple test uniformly more powerful than the Bonferroni correction. It is named after Sture Holm, who codified the method, and Carlo Emilio Bonferroni[4]. The steps to apply it are as follows:

1. Sort all m p-values in increasing order: p(1),p(2),. . . ,p(m).

2. Multiply p(k) by (m-k+1) to get adjusted p-value.

3. If an adjusted p-value p(i)<p(i-1), then we increase it to p(i-1).
4. If a p-value is greater than one, then it is reduced to one.
5. If the adjusted p-values are smaller than α then we reject the null hypothesis.

The results from post hoc Holm's tests are as follows:

1. Heart rate

	$p_{(k)} \times (m-1+k)$	Adjusted p-values
group b vs group c	0.003×3	0.009
group a vs group c	0.006×2	0.013
group a vs group b	0.88×1	0.88

Table 3.6: p-values after applying post hoc Holm's test on heart rate ANOVA.

Hence, change in heart rate when participants from group b are subjected to mental stress is significantly different compared to group c at 5% significance level. Similar can be said for group c and group a.

2. LnSDNN

	$p_{(k)} \times (m-1+k)$	Adjusted p-values
group a vs group c	$1.15 \times 10^{-7} \times 3$	3.43×10^{-7}
group b vs group c	0.0003×2	5.6×10^{-4}
group a vs group b	0.007×1	0.007

Table 3.7: p-values after applying post hoc Holm's test on lnSDNN ANOVA.

Hence, change in LnSDNN is significantly different in each group when participants are subjected to mental stress.

3. LnLF

	$p_{(k)} \times (m-1+k)$	Adjusted p-values
group a vs group c	$4.47 \times 10^{-11} \times 3$	1.34×10^{-10}
group b vs group c	$5.35 \times 10^{-6} \times 2$	1.07×10^{-5}
group a vs group b	0.0007×1	0.0007

Table 3.8: p-values after applying post hoc Holm's test on lnSDNN ANOVA.

Hence, change in LnLF is significantly different in each group when participants are subjected to mental stress.

4. LnHF

	$p_{(k)} \times (m-1+k)$	Adjusted p-values
group a vs group c	$8.48 \times 10^{-7} \times 3$	2.54×10^{-6}
group b vs group c	0.001×2	0.003
group a vs group b	0.008×1	0.008

Table 3.9: p-values after applying post hoc Holm's test on lnHF ANOVA.

Hence, change in LnHF is significantly different in each group when participants are subjected to mental stress.

5. LnLF/HF

	$p_{(k)} \times (m-1+k)$	Adjusted p-values
group a vs group c	0×3	0
group b vs group c	$1.45 \times 10^{-8} \times 2$	2.9×10^{-8}
group a vs group b	$3.23 \times 10^{-6} \times 1$	3.23×10^{-6}

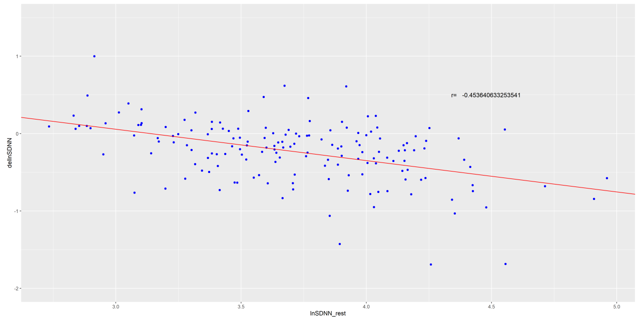
Table 3.10: p-values after applying post hoc Holm's test on lnLF/HF ANOVA.

Hence, change in LnLF/HF is significantly different in each group when participants are subjected to mental stress.

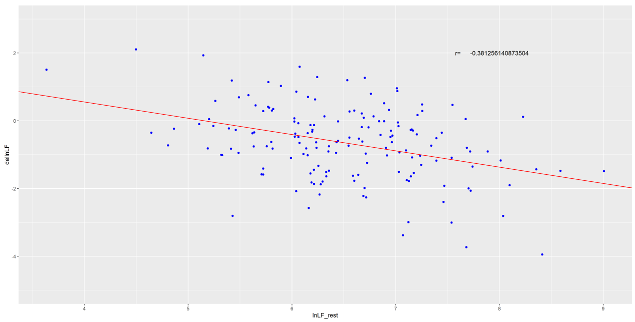
4. Explaining the Observations

The researchers hypothesized that these observations can be explained by RTM. To confirm this, they plotted the graph for change in measurements vs baseline measurements.
The graphs were as follows:

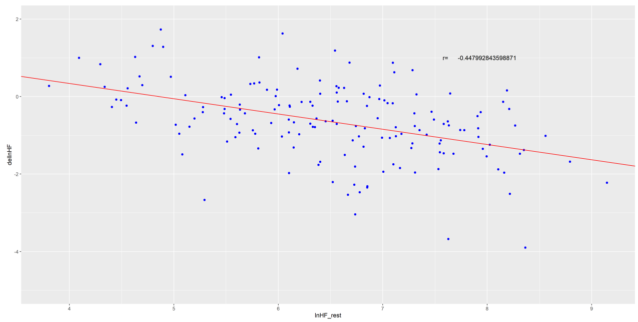
(a) Change in heart rate vs baseline heart rate



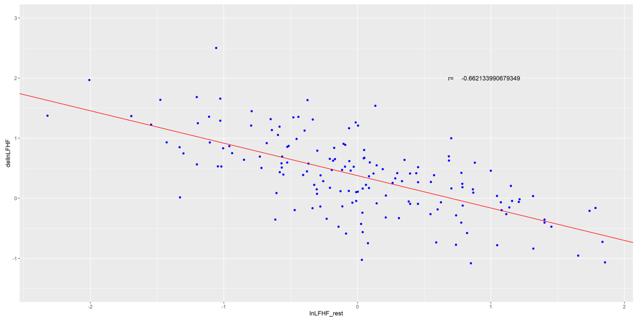
(b) Change in LnSDNN vs baseline LnSDNN



(c) Change in LnLF vs baseline LnLF



(d) Change in LnHF vs baseline LnHF



(e) Change in LnLF/HF vs baseline LnLF/HF

These graphs suggest that there is significant association between baseline levels and effects of mental stress. The negative correlation also favours the hypothesis that RTM plays a role with it. To prove this, we use ANCOVA(Analysis of Covariance).

4.1 ANCOVA

ANCOVA evaluates whether the means of a dependent variable are equal across levels of a categorical independent variable, while statistically controlling for the effects of other continuous variables that are not of primary interest, known as covariates.

ANCOVA is an extension of ANOVA in which main effects and interactions are assessed after dependent variable measurements are adjusted for differences associated with one or more covariates that are measured before the dependent variables and are correlated with it. ANCOVA is a way of controlling for initial individual differences which could not be randomized. We use ANCOVA to determine the effects of independent variable on dependent variable after they are adjusted for the presence of covariance.

4.1.1 Purpose of ANCOVA

1. To increase the sensitivity of the test for main effects and interactions by reducing the error term. The error term is adjusted for and reduced by the relationship between dependent variable and covariance.

2. To adjust the means on the levels of dependent variable itself to what they would be if all subjects had same covariate measurements.

4.1.2 Assumptions of ANCOVA

Before using ANCOVA, we make the following assumptions:

1. Normality

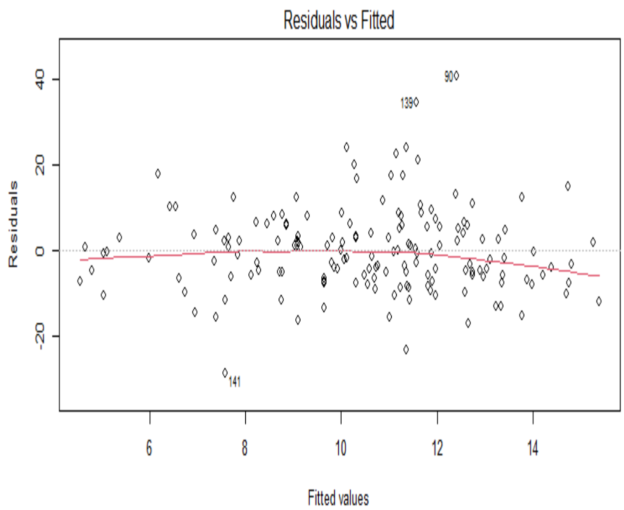
2. Homogeneity of variance

3. Independence

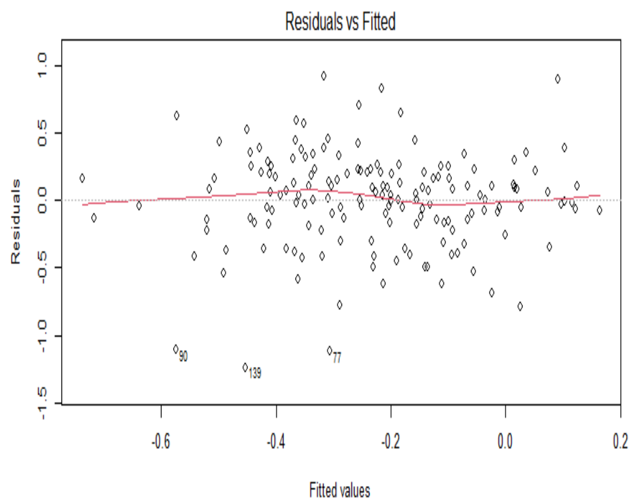
4. Linearity:- For each level of the independent variable, there is a linear relationship between the dependent variable and the covariate.

5. Homogeneity of regression slopes:- The difference between regression slopes for different groups should be insignificant.

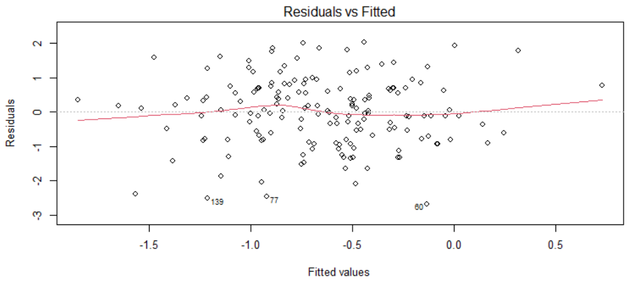
For linearity, we plot a residual vs fitted plot. Residual vs fitted plot is a scatter plot with residuals on y-axis and fitted values on x-axis. If a linear model is suitable for a data set, then its residuals will be more or less randomly distributed around line y=0.



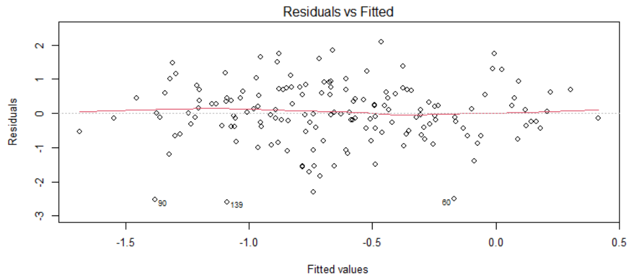
(a) Residuals vs fitted plot for heart rate



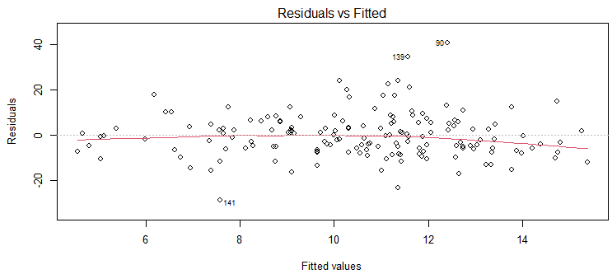
(b) Residuals vs fitted plot for LnSDNN



(c) Residuals vs fitted plot for LnLF



(d) Residuals vs fitted plot for LnHF



(e) Residuals vs fitted plot for LnLF/HF

Thus, the assumption of linearity is not violated.

For homogeneity of regression slopes,we first calculate the slopes with and without the interaction term. If the slopes are significantly different then ANCOVA should not be performed. After running some R-code, we see that the slopes are insignificantly different in all the measurements. Thus, we can safely apply ANCOVA.

For ANCOVA, the dependent variable , covariant and independent variable taken are change in measurements during stress, baseline measurements and group respectively. We created a linear model using lm() function in R with an interaction term. Using R, we performed ANCOVA which tests the null hypothesis that mean change is insignificantly different among the three groups after adjusting for baseline measurements.

5. Conclusion

Examining the results suggested that mean change in HRV is insignificantly different among the three different groups after adjusting for the baseline HRV.

This suggests that if everyone had similar HRV before the stress, then their HRV would have been similar during stress as well.

Thus, we can safely conclude that for extreme observations, RTM was the major cause for change in HRV during mental stress.

6. References

1. Regression to the mean,wikipedia.

2. HRV,wikipedia

3. Shaffer, Fred, Rollin McCraty, and Christopher L. Zerr. "A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability." Frontiers in psychology 5 (2014): 1040. Heart rate,wikipedia

4. post hoc holm's test,wikipedia

5. C.E. Davis, The effect of regression to the mean in epidemiologic and clinical studies, Am. J. Epidemiol. 104 (1976) 493–498.

6. P. Castiglioni, G. Parati, A. Civijian, L. Quintin, M. Di Rienzo, Local scale exponents of blood pressure and heart rate variability by detrended fluctuation analysis: effects of posture, exercise, and aging, IEEE Trans. Biomed. Eng. 56 (2009) 675–684.

7. M.P. Tarvainen, J.P. Niskanen, J.A. Lipponen, P.O. Ranta-Aho, P.A. Karjalainen, Kubios HRV–heart rate variability analysis software, Comput. Methods Prog. Biomed. 113 (2014) 210–220.

8. A.G. Barnett, J.C. Van Der Pols, A.J. Dobson, Regression to the mean: what it is and how to deal with it, Int. J. Epidemiol. 34 (2004) 215–220.