Report 1

Daniel J. Naumenko

8 February 2022

Report Introduction

Rebecca Brittain and Caitlin O'Connell have increased the size of the oxidative stress and inflammation biomarker samples, and cleaned nutritional data are primarily available. Therefore, I am creating a new first report to detail the exploratory data analysis and eventual hypothesis testing for the caloric and dietary restriction and aging paper.

I have made the CRediT Taxonomy of Author Contributions into a table, which is presented in it's preliminary form below. I have gone through and identified where I think people belong based on memory (note that I will not forget middle initials when it comes time for publication). Please let me know if the author list should change, and if specific contributions should be added or removed. This is open and subject to change, pending discussion among authors. As of now, the key points for inclusion are that: Rebecca, Caitlin, and Erin have done lab work after I left the lab, and Tim is included in this if cortisol is used; Tim and Rebecca worked on nutritional calculations; Suci and Erin supervised/managed the field site, the lab, etc; and I conceptualized the projects revolving around oxidative stress and aging in orangutans, dietary restriction, etc. Blank spaces have yet to be filled in.

As of now, I think I was the only one to conceive of the caloric restriction project in its form, given I introduced the biomarkers into the lab and explored caloric restriction as having a causal role in lifespan and orangutan health.

Exploratory Data Analysis

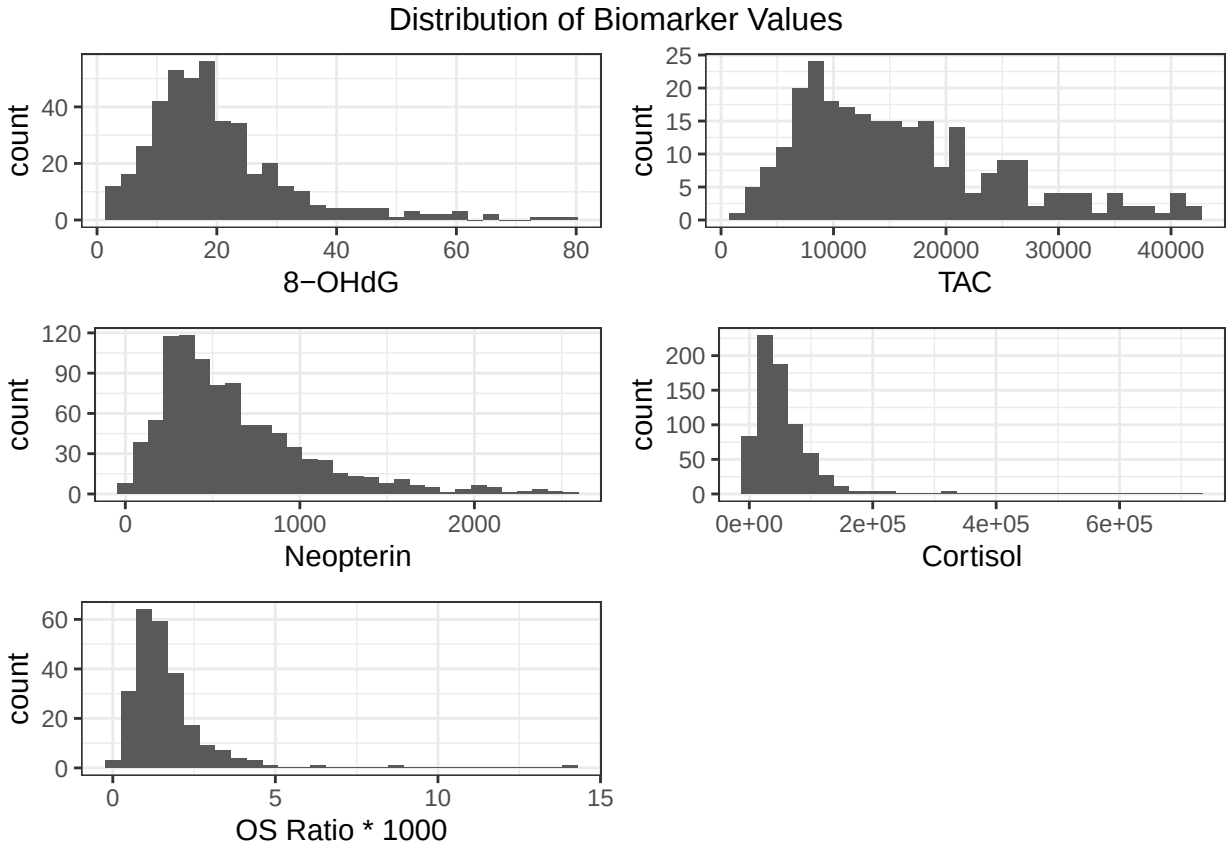
All 4 biomarkers have been screened for outliers and other exclusionary factors. This first pass INCLUDES non-first-morning voids. It will be repeated with those removed to see if results significantly change. As this is a first pass prior to the remainder of samples being run, I decided to include everything for now. Samples with a CV $> 15\%$ were removed. Samples with a specific gravity $<= 1.003$ were also removed as being too dilute. Finally, for each biomarker, those observations which exceeded +3SD or -3SD were also removed. I am debating on whether this step is appropriate. It may be relevant to include for 8-OHdG, TAC, or neopterin, but I think the outlier procedure should remain in place for cortisol. This biomarker has some extreme values that will almost certainly impact data analysis. See distributions below under Variation.

CRediT Taxonomy Category	Description
Conceptualization	Ideas; formulation or evolution of overarching research goals and aims.
Data Curation	Management activities to annotate (produce metadata), scrub data and maintain research
Formal Analysis	Application of statistical, mathematical, computational, or other formal techniques to ana
Funding Acquisition	Acquisition of the financial support for the project leading to this publication.
Investigation	Conducting a research and investigation process, specifically performing the experiments,
Methodology	Development or design of methodology; creation of models.
Project Administration	Management and coordination responsibility for the research activity planning and execut
Resources	Provision of study materials, reagents, materials, patients, laboratory samples, animals, in
Software	Programming, software development; designing computer programs; implementation of th
Supervision	Oversight and leadership responsibility for the research activity planning and execution, in
Validation	Verification, whether as a part of the activity or separate, of the overall replication/reproc
Visualization	Preparation, creation and/or presentation of the published work, specifically visualization
Writing - Original Draft	Preparation, creation and/or presentation of the published work, specifically writing the in
Writing - Review and Editing	Preparation, creation and/or presentation of the published work by those from the origina

Note:

For more information: https://casrai.org/credit/

Variation



Unsurprisingly, all biomarkers are right-skewed. Even with outliers removed following the protocol described above, cortisol's distribution is the most extreme, with a poor spread of extremely high values. Compare this with 8-OHdG and neopterin which also have several high values but a more representative spread of those high values. Also of note, TAC has a much wider and more evenly spread distribution of values relative to 8-OHdG. The purpose of the OS Ratio (which is 8-OHdG / TAC * 1000) is to help capture the relative impact

Table 1: Descriptive Statistics: 8-OHdG

age_class_simp	n_obs	mean	median	mode_whole	sd	minimum	maximum	IQR	
Adult Female	214	19.65749	17.05924	15	11.548525	1.347310	77.60035	10.531125	12.583
Adult Male	179	21.17877	18.72750	23	13.517343	1.860571	77.66000	14.397134	12.253
Inf-Juv	26	15.72681	13.93443	11	6.291791	6.540000	28.94182	9.574097	10.636

Table 2: Descriptive Statistiscs: Total Antioxidant Capacity (TAC)

age_class_simp	# of Observations	mean	median	mode_whole	sd	minimum	maximum	I
Adult Female	141	16421.31	14982.928	12461	8868.500	1280.565	41910.11	11629.
Adult Male	103	16366.79	14367.454	6176	9645.253	2881.975	41817.24	14910.
Inf-Juv	16	11220.95	9479.229	17309	5383.736	3362.304	20987.28	7190.

of each biomarker. Oxidative stress is the inability for TAC to stop oxidative damage, meaning that levels of 8-OHdG must be interpreted in light of TAC values. A low 8-OHdG value with a high TAC value indicates that something is increasing oxidative stress but that antioxidant defenses are currently sufficient to handle the issue. If both values are high, then the antioxidant defense are responding but being overwhelmed.

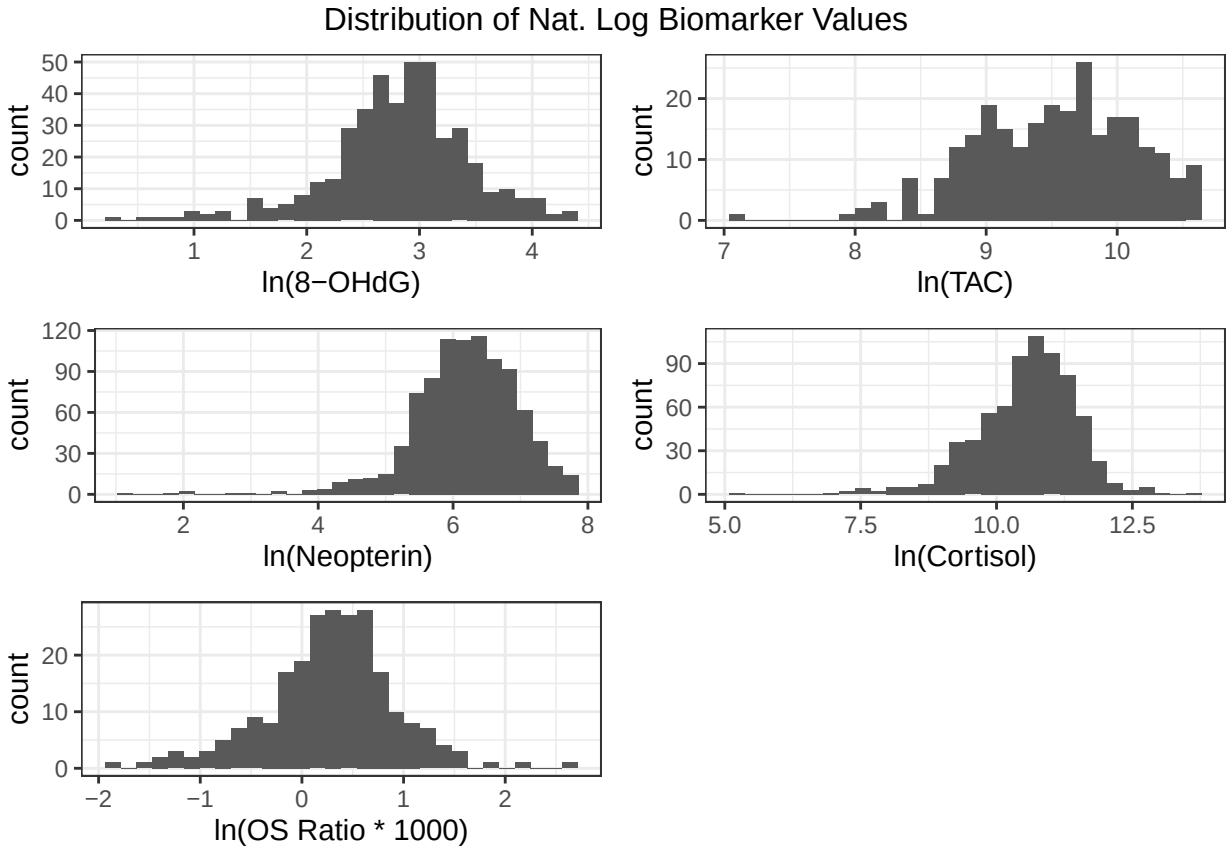


Table 3: Descriptive Statistiscs: Oxidative Stress Ratio (x1000)

age_class_simp	# of Observations	mean	median	mode_whole	sd	minimum	maximum	I
Adult Female	128	1.545967	1.272405	1	1.5166332	0.1584740	14.215600	0.9213
Adult Male	96	1.703229	1.489659	1	1.0580669	0.2354285	6.075596	1.0538
Inf-Juv	15	1.844847	1.759819	2	0.9621119	0.6085636	4.156674	1.1587

Table 4: Descriptive Statistics: Neopterin

age_class_simp	# of Observations	mean	median	mode_whole	sd	minimum	maximum	IQ
Adult Female	417	580.5429	455.4160	319	427.0455	8.226857	2492.338	485.645
Adult Male	444	670.4300	537.7356	425	458.2127	3.456545	2554.969	501.436
Inf-Juv	65	604.2654	529.8305	365	380.8399	64.670400	2102.397	407.122

The mode for this table is not very interpretable. To calculate mode, I rounded to whole numbers for consistency. It is entirely unsurprising that 1 is the mode for both adult males and females, as it is the center of a very tiny range.

Covariation

I should note that it might be valuable to look at creatinine corrections because all available literature corrects urinary 8-OHdG with creatinine. The only difference was Nicole Thompson-Gonzalez's study out of Melissa Emery Thompson's lab. However, she did not provide any descriptive statistics, only model outputs. Therefore I do not know what sort of range her data had (which is rather frustrating). They included some plots, but . . . .*sigh*. It looks like, for 8-OHdG, the median for before, during, and after their so called epidemic is 8, 10, and 12 ng/mL, with almost all points falling below 20 ng/mL. They put in a few other graphs. However, they did not adjust oxidative damage for antioxidant capacity at all, despite also measuring TAC. I have included Supplementary Tables 5 and 6 from this publication below.

Now curiously, she found no correlation between 8-OHdG and TAC. If you see the correlation plot below the two tables, there is a reasonably strong spearman correlation between 8-OHdG and TAC in our data. From just a spearman, there is no correlation between 8-OHdG and neopterin. These are, of course, raw biomarkers

Table 5: Descriptive Statistics: Cortisol

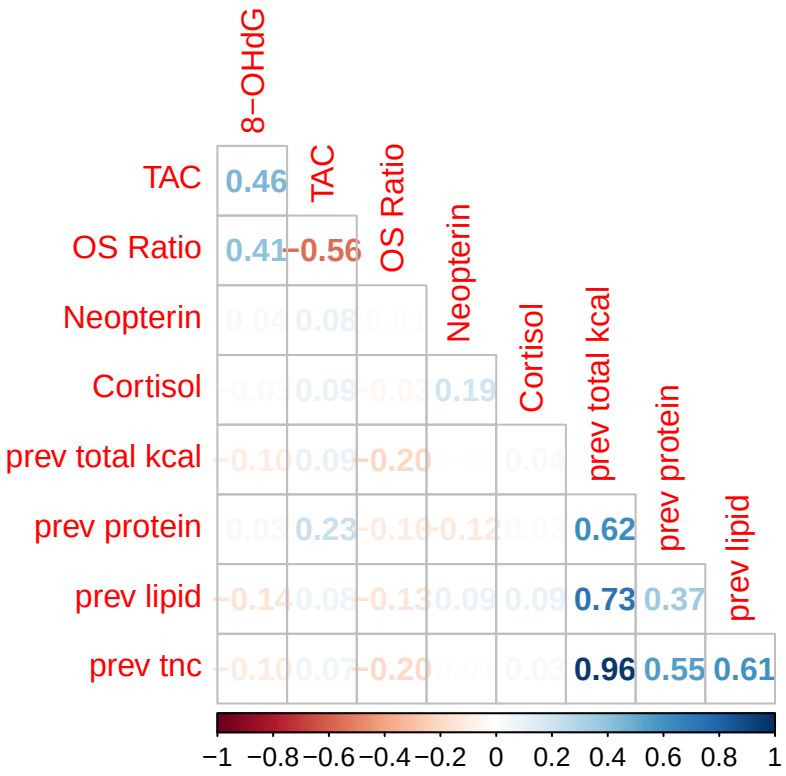
age_class_simp	# of Observations	mean	median	mode_whole	sd	minimum	maximum	IQ
Adult Female	347	48054.08	37084.94	45290	49952.26	160.5823	533304.8	40959.7
Adult Male	317	63540.71	51399.48	54195	60894.92	1055.6400	721880.8	54012.6
Inf-Juv	51	41828.75	31154.93	209544	37777.26	7996.5273	209544.2	28680.5

Table 6: Table S5: Correlations between OS biomarkers

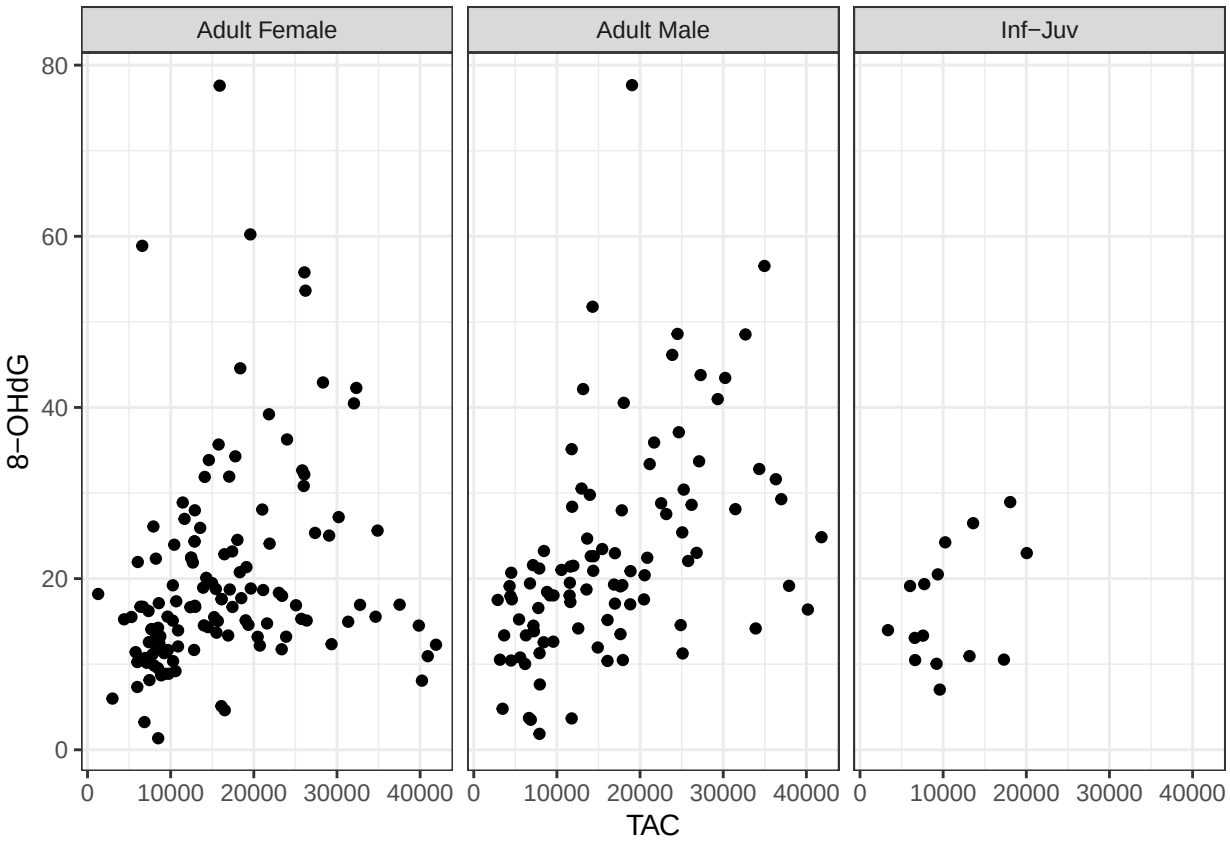
Biomarker 1	Biomarker 2	n_individual	n_samles	rho	p-value
8-OHdG					
8-OHdG	**Isop**	**26**	**336**	**0.29**	**<0.001**
8-OHdG	**MDA-TBARS**	**33**	**484**	**0.10**	**0.03**
8-OHdG	**Neo**	**37**	**631**	**0.09**	**0.03**
8-OHdG	TAC	32	470	-0.05	0.3
Isop					
Isop	**MDA-TBARS**	**26**	**291**	**0.12**	**0.047**
Isop	Neo	26	300	0.07	0.23
Isop	**TAC**	**26**	**327**	**0.49**	**<0.001**
MDA-TBARS					
MDA-TBARS	Neo	33	381	0.06	0.27
MDA-TBARS	TAC	26	288	0.00	0.98
Neo					
Neo	**TAC**	**32**	**441**	**0.13**	**0.01**

Table 7: Table S6: Cross-sectional variation in OS biomarker by individual age and sex

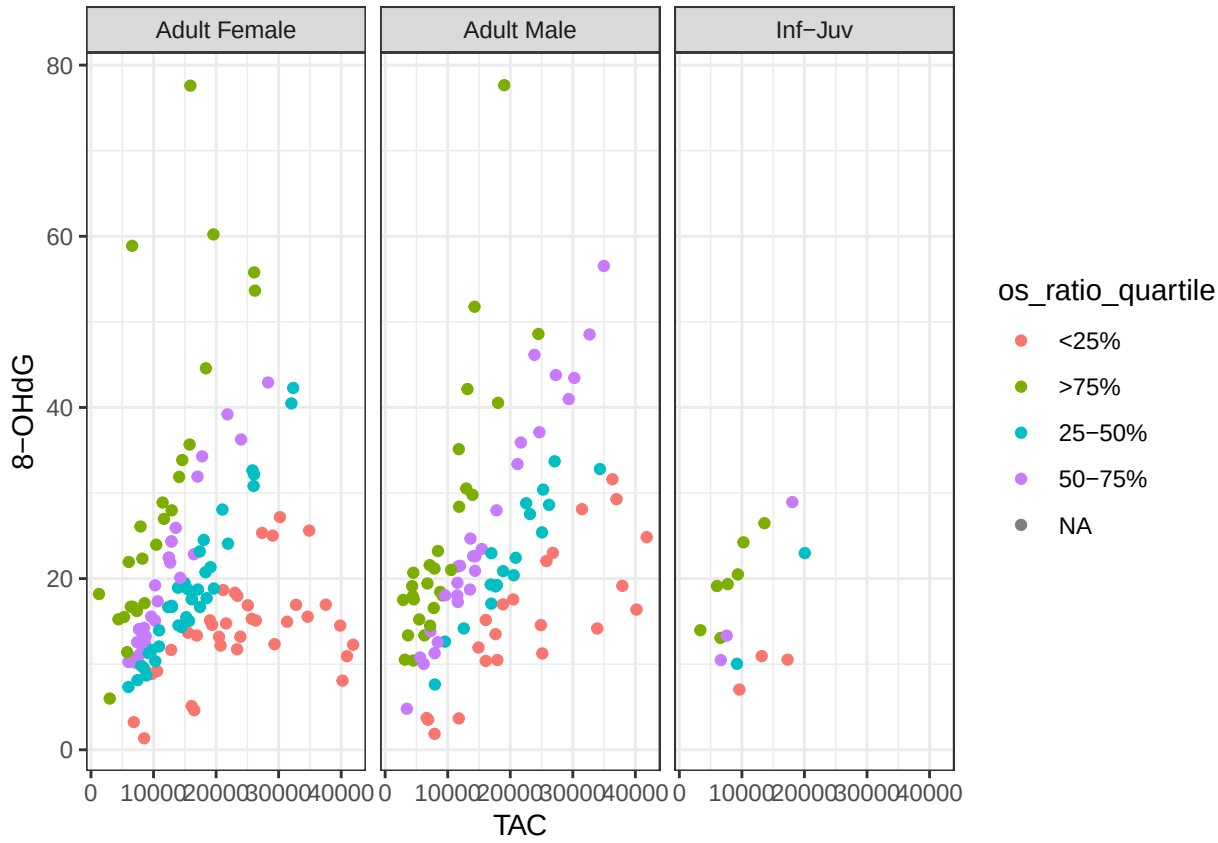
Biomarker	n_individuals	n_samples	predictor	beta	SE	CI_95	p	percentRE
8-OHdG								
8-OHdG	**36**	**582**	**Intercept**	**2.82**	**0.06**	**2.71 - 2.94**	**< 0.001**	**6.92**
8-OHdG	36	582	sex (M)	-0.07	0.08	-0.32	0.401	6.92
8-OHdG	36	582	age	-0.10	0.05	-0.18	0.028	6.92
Isoprostanes								
Isoprostanes	**23**	**182**	**Intercept**	**1.34**	**0.13**	**1.08 - 1.6**	**< 0.001**	**18.46**
Isoprostanes	23	182	sex (M)	-0.05	0.19	-0.76	0.793	18.46
Isoprostanes	23	182	age	0.05	0.10	-0.38	0.606	18.46
MDA-TBARS								
MDA-TBARS	**31**	**249**	**Intercept**	**2.59**	**0.07**	**2.45 - 2.72**	**< 0.001**	**6.53**
MDA-TBARS	31	249	sex (M)	-0.04	0.09	-0.35	0.692	6.53
MDA-TBARS	31	249	age	-0.01	0.04	-0.16	0.829	6.53
Neopterin								
Neopterin	**36**	**409**	**Intercept**	**6.97**	**0.06**	**6.84 - 7.09**	**< 0.001**	**8.00**
Neopterin	36	409	sex (M)	0.05	0.09	-0.36	0.562	8.00
Neopterin	36	409	age	-0.01	0.05	-0.18	0.831	8.00
TAC								
TAC	31	295	Intercept	0.10	0.31	-1.2	0.741	14.97
TAC	31	295	sex (M)	-0.32	0.47	-1.84	0.504	14.97
TAC	31	295	age	-0.02	0.24	-0.92	0.927	NA



These correlations here are from Tuanan data, while the two tables above are from Nicole's paper on chimps.



The direction of relationships seems to be the same, roughly speaking, across all 3 age-sex classes.

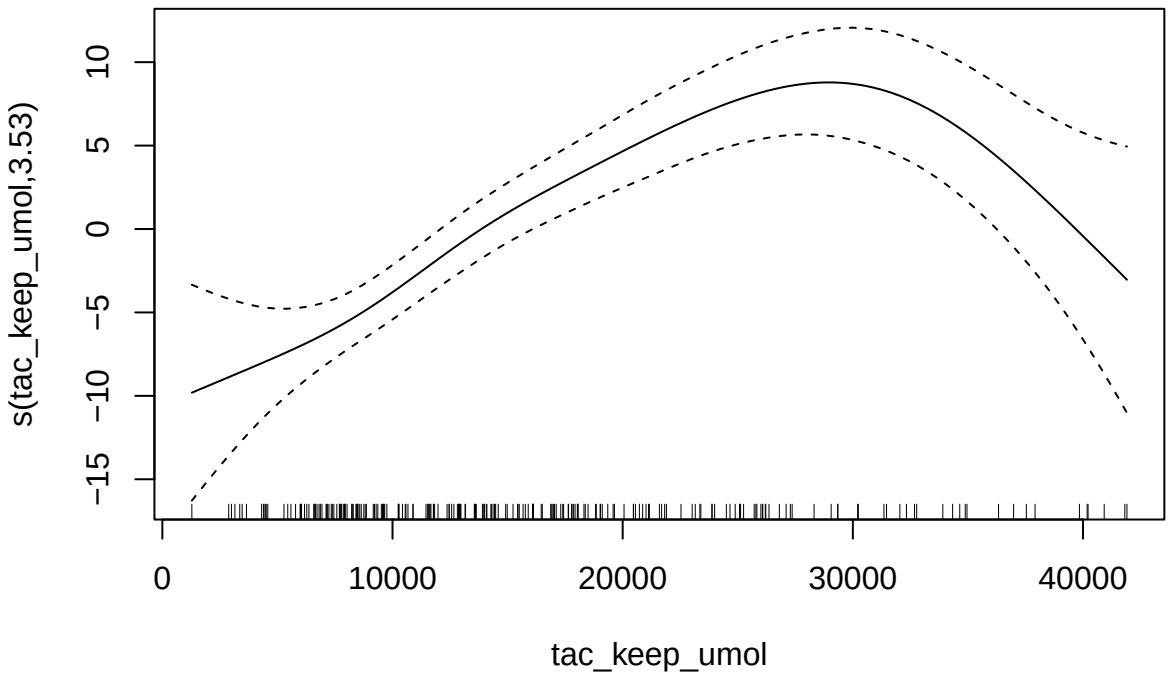


Caloric Restriction and Oxidative Stress

I need to make sure the list of orangutan IDs is correct, as well as double check that age-sex classes are input properly.

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## ohdg_keep ~ s(tac_keep_umol)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  20.9062     0.6943   30.11   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                     edf Ref.df     F p-value
## s(tac_keep_umol) 3.532  3.532 17.42  <2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.205
##   Scale est. = 114.72    n = 239
```



So the two biomarkers increase with each other. The way I am thinking of this is antioxidant defenses are increasing in response to free radicals, but not necessarily dealing with all them, thus resulting in damage. At the upper end of TAC, two things could be happening... 1) there is just insufficient samples. The rug plot seems to show a spread of data so I'm not sure this is necessarily the case. The second option is that so many antioxidants are being produced that free radicals are being dealt with. Another option is that there is a large influx of dietary antioxidants that have nothing to do with free radicals, or that antioxidants are increasing for some other reason.

Make sure to assess the fit.