

# Statistical study of backpain

*s1899215*

*February 10, 2019*

## 1 Introduction

## 2 Exploring Raw data

### 2.1 Biomarker data

Table 1: Biomarker raw data

Biomarker	IL-8	VEGF-A	OPG	TGF-beta-1	IL-6	CXCL9	CXCL1	IL-18	CSF-1
126-0weeks	7.63	11.51	10.20	8.83	3.52	6.16	9.45	7.91	8.41
126-6weeks	7.12	11.59	10.41	8.87	3.89	6.12	9.06	7.92	8.39
127-0weeks	6.93	10.92	10.30	6.59	2.73	6.14	7.31	7.95	8.40
127-6weeks	7.16	11.58	10.39	8.61	2.60	6.35	8.61	7.94	8.51
127-12months	6.87	11.13	10.25	7.44	3.92	6.15	8.79	7.94	8.46

Table 2: NAs in Biomarker data

Biomarker	IL-8	VEGF-A	OPG	TGF-beta-1	IL-6	CXCL9	CXCL1	IL-18	CSF-1
	4	3	4	4	4	4	4	4	4

From the Biomarker data we observed the following:

Number of observations: 351

Number of columns (proteins tested): 10

4 rows contain NA values in the raw data. These rows are removed from the data, as the NAs can cause some errors in our calculation.

### 2.2 Patient data

Table 3: Covariates raw data

PatientID	Age	Sex (1=male, 2=female)	Smoker (1=yes, 2=no)	VAS-at-inclusion	Vas-12months
1	56	1	2	3.0	4.0
3	32	1	2	7.2	0.5
4	43	2	2	2.7	0.5
5	25	2	2	3.0	3.9
6	39	1	2	3.5	5.0

Table 4: NAs in Covariate data

PatientID	Age	Sex (1=male, 2=female)	Smoker (1=yes, 2=no)	VAS-at-inclusion	Vas-12months
0	0	0	0	0	2

From the covariate data set we observed the following:

Number of observations: 118

Number of columns: 6

2 number of rows removed from the original data

### 2.2.1 Patient data - simple analysis

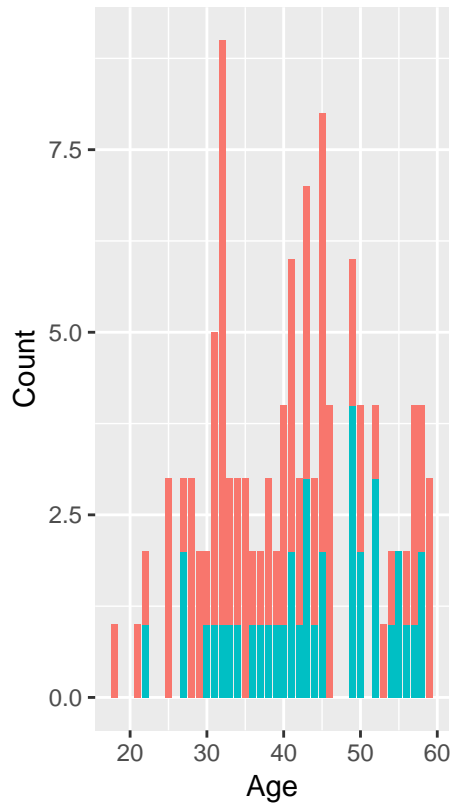


Fig 1: number of smokers in each Age group

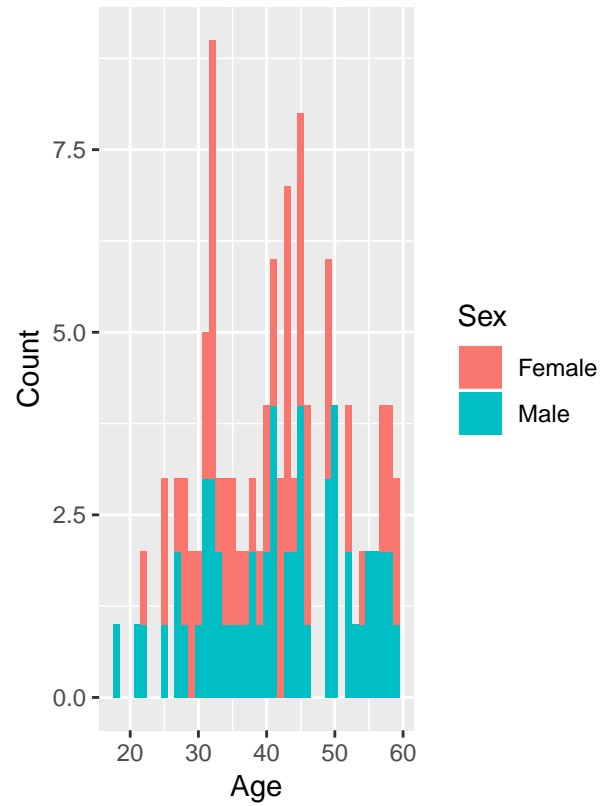


Fig 2: Different sex in each age group

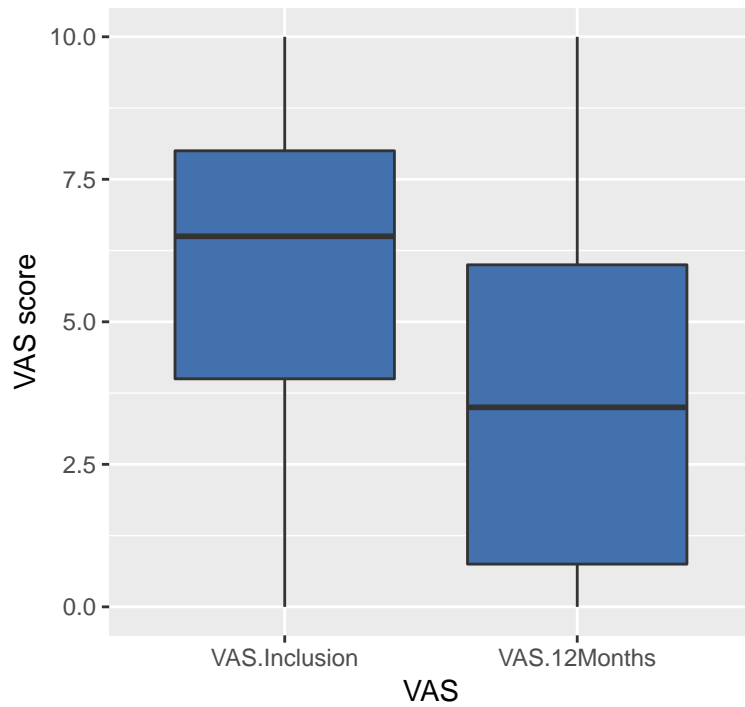


Fig 3: Comparing VAS at inclusion and at 12 months

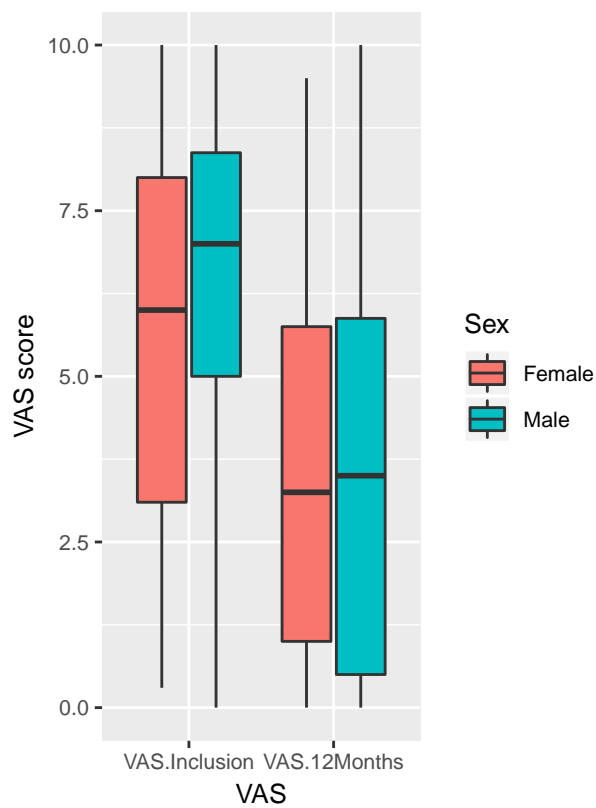


Fig 4: Comparing VAS on different sex

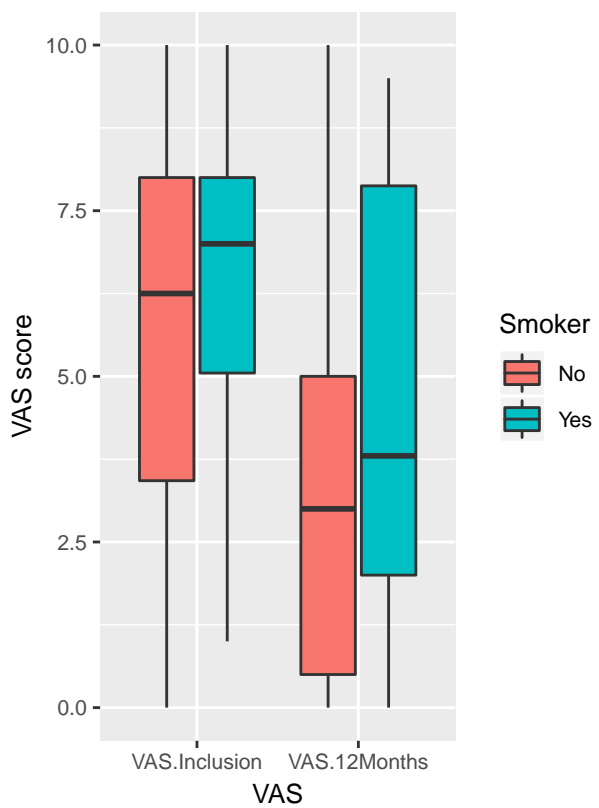


Fig 5: Comparing VAS on smokers

### 3 Cleaning/Manipulating Data

Table 5: Biomarker data after separating the patient ID and week of observation

PatientID	WeeksObs	IL-8	VEGF-A	OPG	TGF-beta-1	IL-6	CXCL9	CXCL1	IL-18	CSF-1
126	0	7.63	11.51	10.20	8.83	3.52	6.16	9.45	7.91	8.41
126	6	7.12	11.59	10.41	8.87	3.89	6.12	9.06	7.92	8.39
127	0	6.93	10.92	10.30	6.59	2.73	6.14	7.31	7.95	8.40
127	6	7.16	11.58	10.39	8.61	2.60	6.35	8.61	7.94	8.51
127	52	6.87	11.13	10.25	7.44	3.92	6.15	8.79	7.94	8.46
128	0	8.62	12.51	10.56	8.51	3.71	7.34	9.90	8.72	8.72
128	6	6.94	11.50	10.51	7.46	3.84	7.14	8.57	8.62	8.51
128	52	6.47	11.05	10.14	6.45	4.65	8.00	8.18	8.71	8.56
129	0	8.16	11.16	10.61	8.76	3.85	5.81	9.18	7.49	8.39
129	6	6.57	10.72	10.23	6.82	2.98	6.11	6.69	7.23	8.16

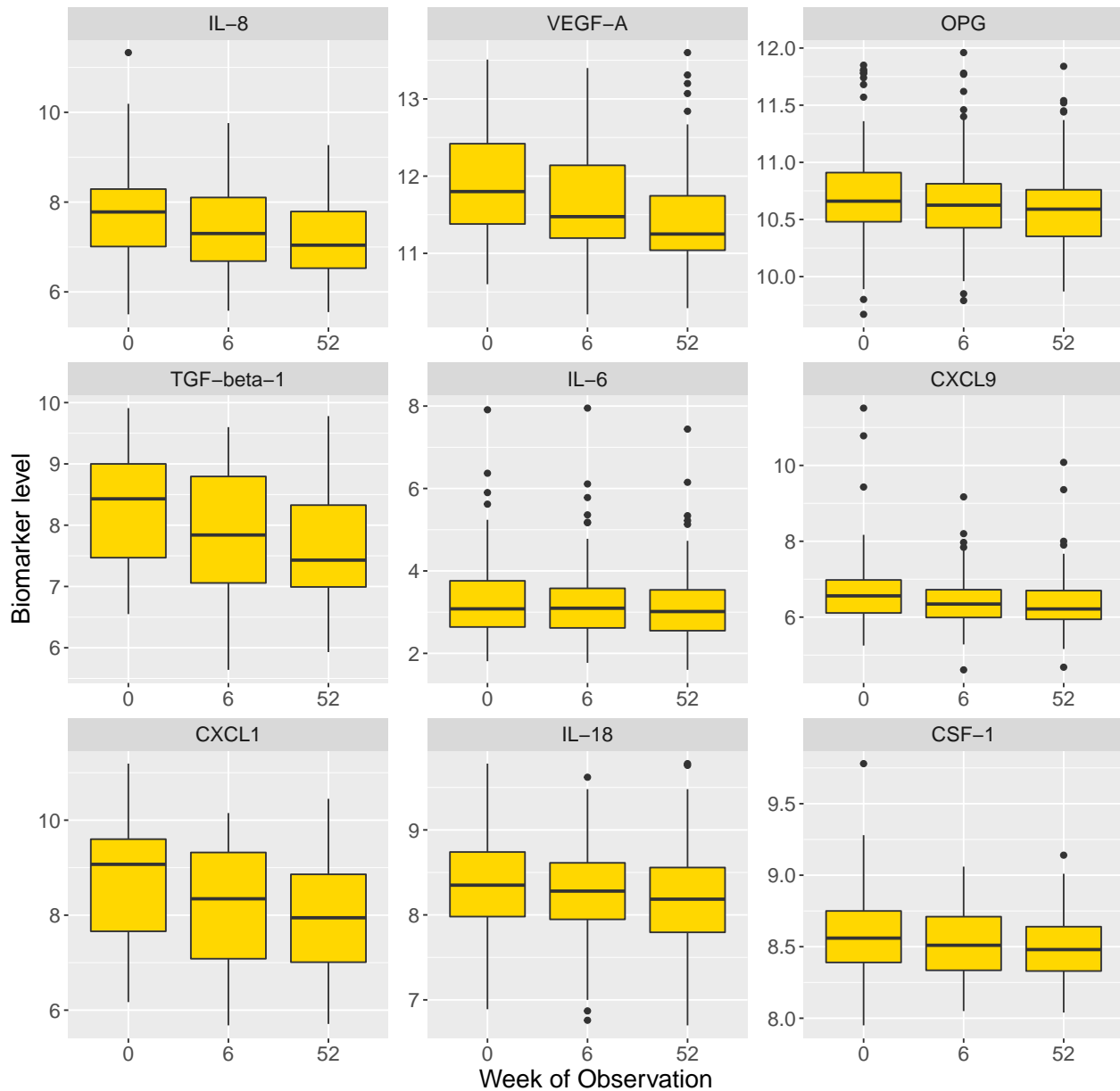


Fig 6: Biomarker value over the course of week of observation

Reference for scaling

## Question 1:

Choose a question and explain why you chose that

Among the patients that showed improvement, did their biomarker level change at each week of observation. If it did, we may be able to utilize biomarker level as a variable to track the progress of pain relief.

Total of 79 will be used for the hypothesis testing. These patients showed improvement for the medication and have biomarker level measured at all three occasion.

## Question 2:

Hypothesis:  $\mu$  is not the same for each week  
Parameters: Random variables: Levels of IL-8, VEGF-A, OPG, TGF-beta-1, IL-6, CXCL9, CXCL1, IL-18, CSF-1  
Distribution: t-distribution since the population  $\mu$  is unknown  
 $n_1 \neq n_2$   
#Question 3: Conduct your hypothesis test  
Paired sample t-test to check the  $\mu$   
Explain why paired

**Week0 == week6**

**Week6 == week52**

**Week0 == Week52**

Table 6: p.vales for each hypothesis test

Biomarker	Week 0 and Week 6	Week 6 and Week 52	Week 0 and Week 52
IL-8	0.0145484	0.0290526	0.0000110
VEGF-A	0.0847731	0.0028904	0.0000034
OPG	0.3927919	0.1302943	0.0234495
TGF-beta-1	0.0328849	0.0617845	0.0000483
IL-6	0.9194881	0.3405674	0.3552379
CXCL9	0.0731432	0.2751667	0.0149160
CXCL1	0.0161861	0.1431040	0.0000329
IL-18	0.2004710	0.2857632	0.0209696
CSF-1	0.0742271	0.0883941	0.0021744

Table 7: Acception of p.values

Biomarker	Week 0 and Week 6	Week 6 and Week 52	Week 0 and Week 52
IL-8	Reject	Reject	Reject
VEGF-A	Accept	Reject	Reject
OPG	Accept	Accept	Reject
TGF-beta-1	Reject	Accept	Reject
IL-6	Accept	Accept	Accept
CXCL9	Accept	Accept	Reject
CXCL1	Reject	Accept	Reject
IL-18	Accept	Accept	Reject
CSF-1	Accept	Accept	Reject

## Interpretation of the results from the table

### Conclusion

#### Q3 a

Potential problem of multiple testing Calculate the probability of making at least one type I error assuming that your tests are independent and that all null hypotheses are true.

#### Q3 b

Bonferroni correction: What it is used for? How to use it Redo the test using this method