Statistical study of backpain

s1899215

February 10, 2019

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 raw data

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

From the Biomarker data we observed the following:

Number of observations: 351

Number of columns (proteins tested): 10

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

The Biomaker column contains both patientID and the time the test was conducted. Therefore this needs to be extracted and registered in seperate columns.

2.2 Patient data

Table 3: covariates raw data

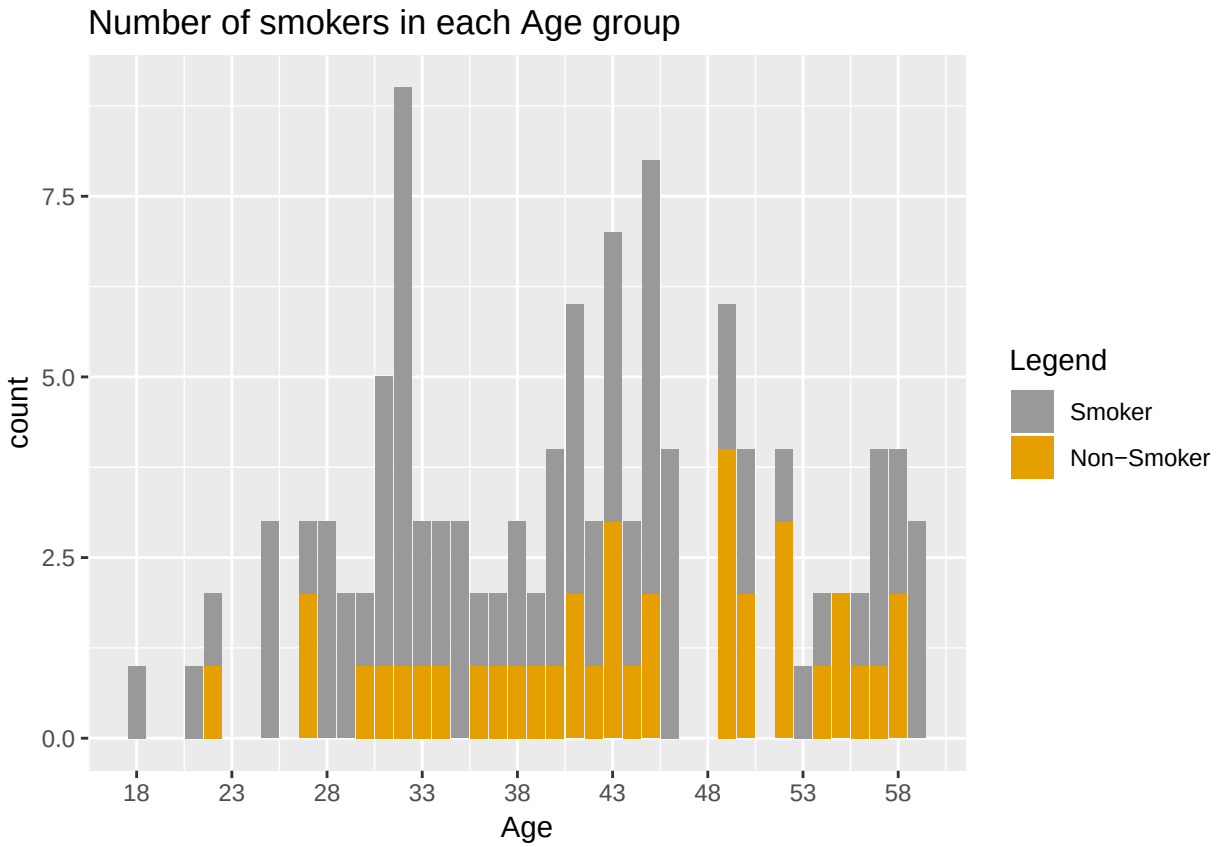
PatientID	Age	Sex (1=male, 2=female)	Smoker (1=yes, 2=no)	VAS-at-inclusion	Vas-12months
1	56	1	2	3.0	4.0
3	32	1	2	7.2	0.5
4	43	2	2	2.7	0.5
5	25	2	2	3.0	3.9
6	39	1	2	3.5	5.0

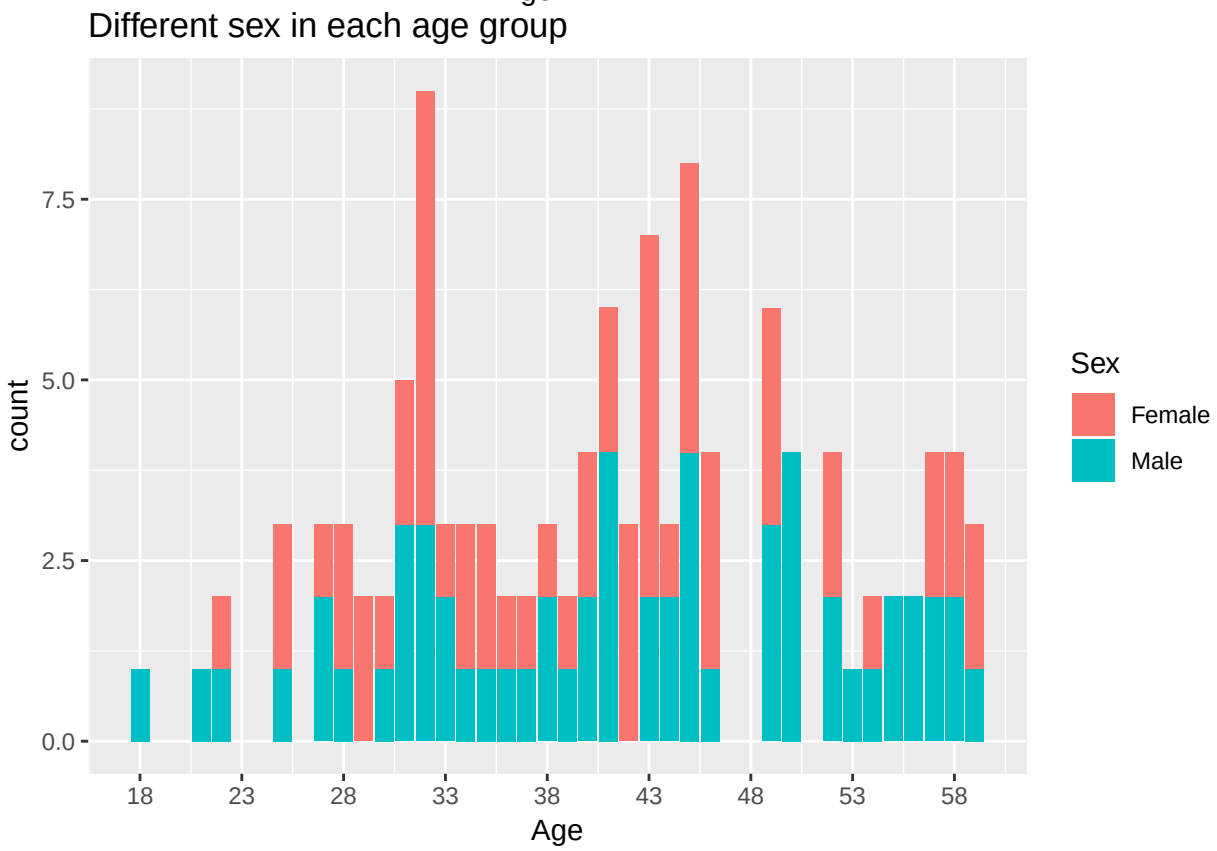
Table 4: NAs in covariates raw data

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

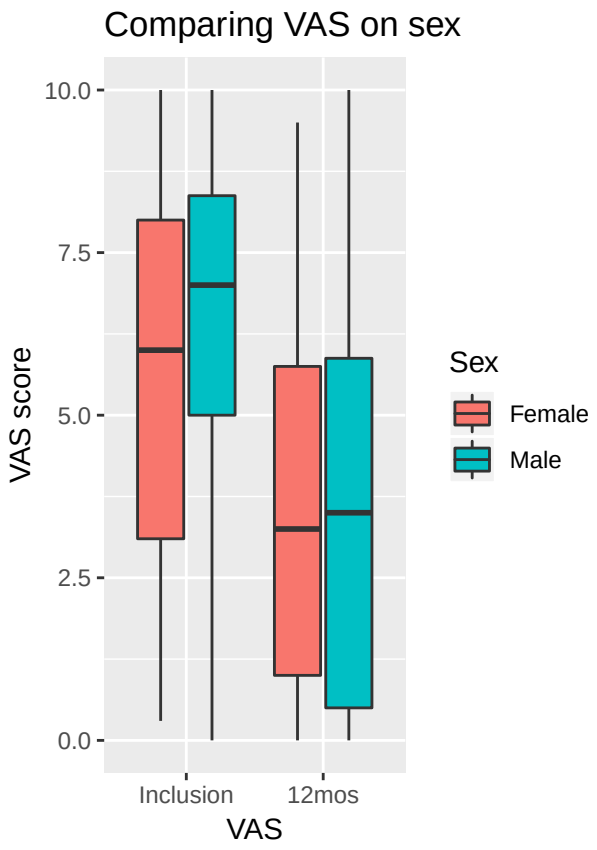
Number of observations: 118 Number of columns (proteins tested): 6 2 patients data are incomplete. Therefore this is removed from future calculation to avoid error or misrepresentation.

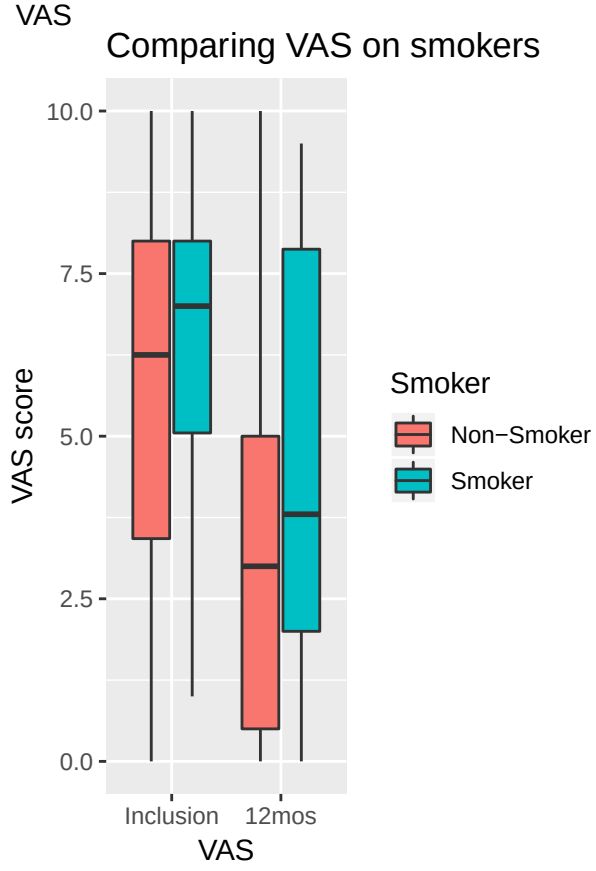
2.2.1 Patient data - simple analysis





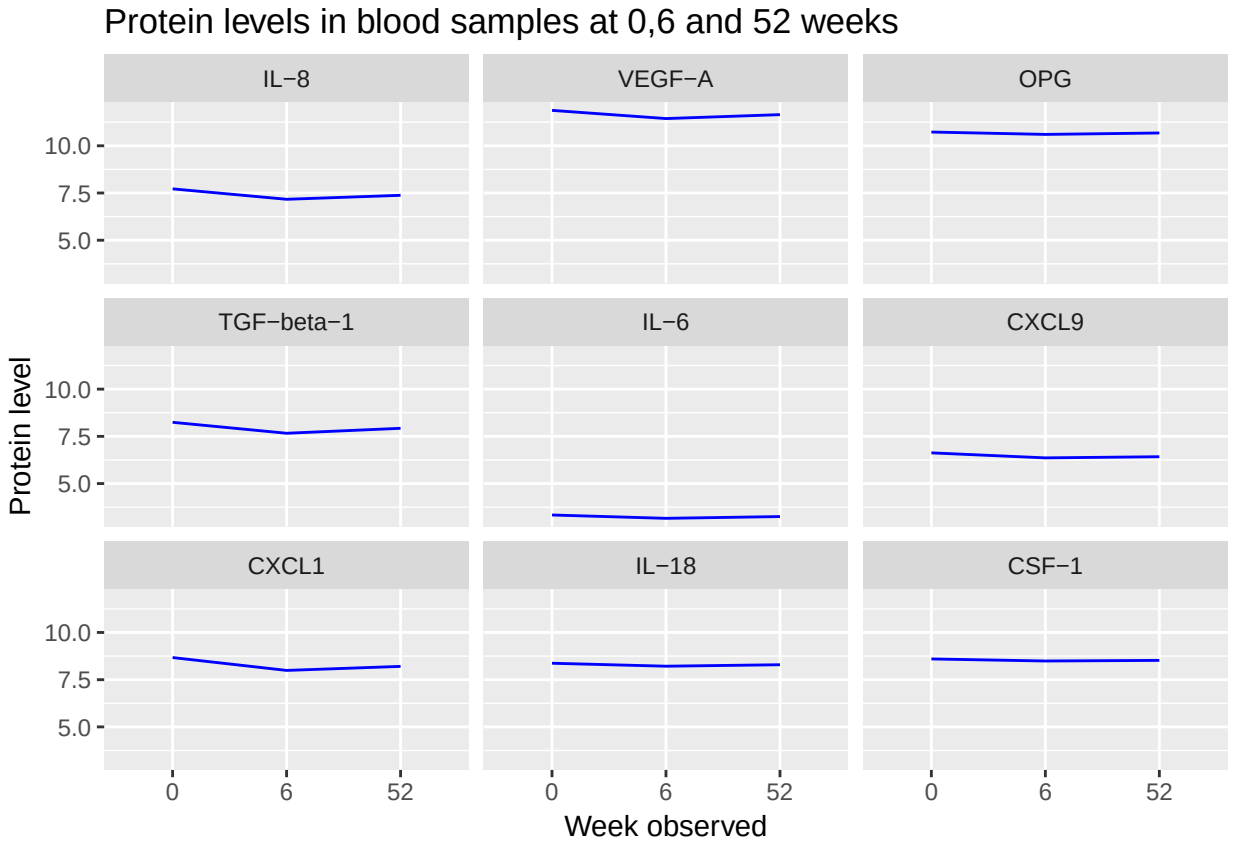

Comparing VAS at inclusion and at 12 months
VAS score
10.0
7.5
5.0
2.5
0.0
Inclusion
12mos
VAS


Comparing VAS on sex
VAS score
10.0
7.5
5.0
2.5
0.0
Inclusion
12mos
VAS
Sex
Female
Male

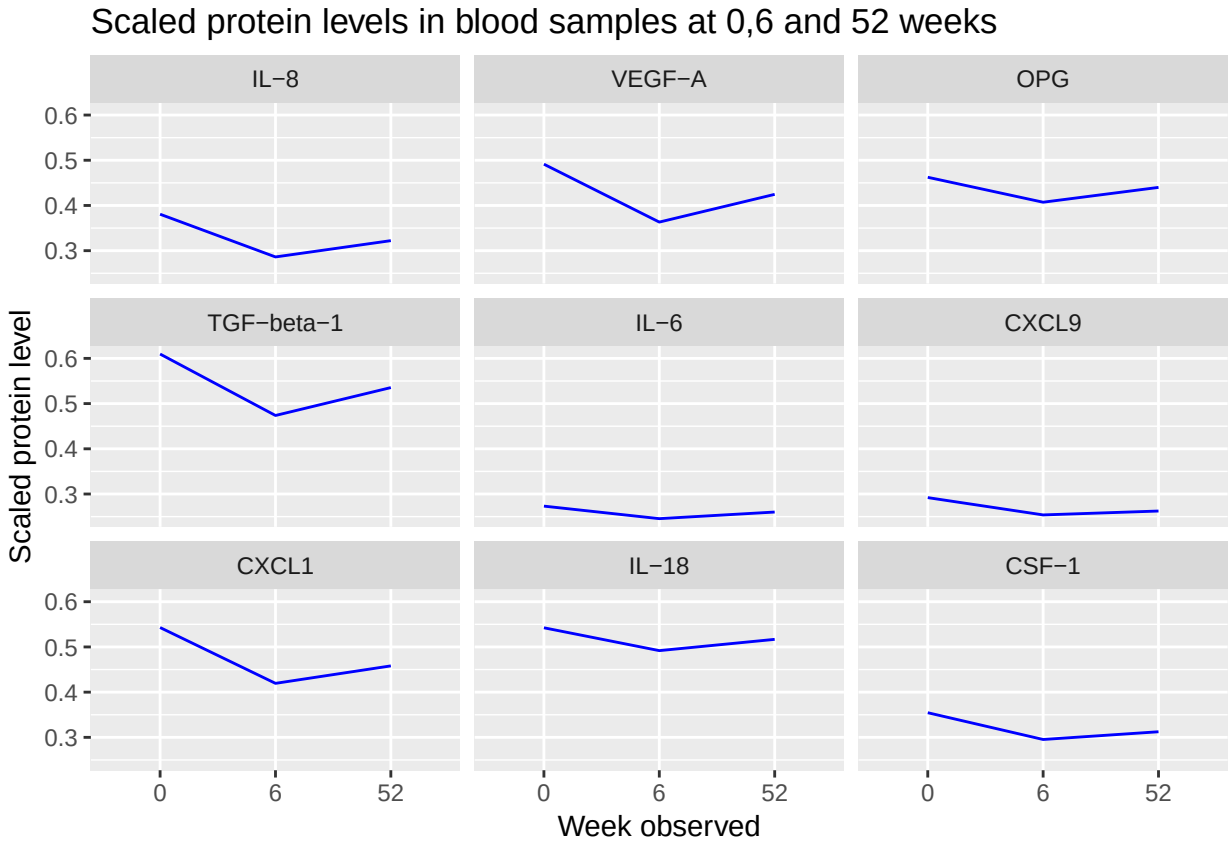

Comparing VAS on smokers
VAS score
10.0
7.5
5.0
2.5
0.0
Inclusion
12mos
VAS
Smoker
Non−Smoker
Smoker

3 Cleaning/Manipulating Data

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




```
## Saving 6.5 x 4.5 in image
```




```
## Saving 6.5 x 4.5 in image
```

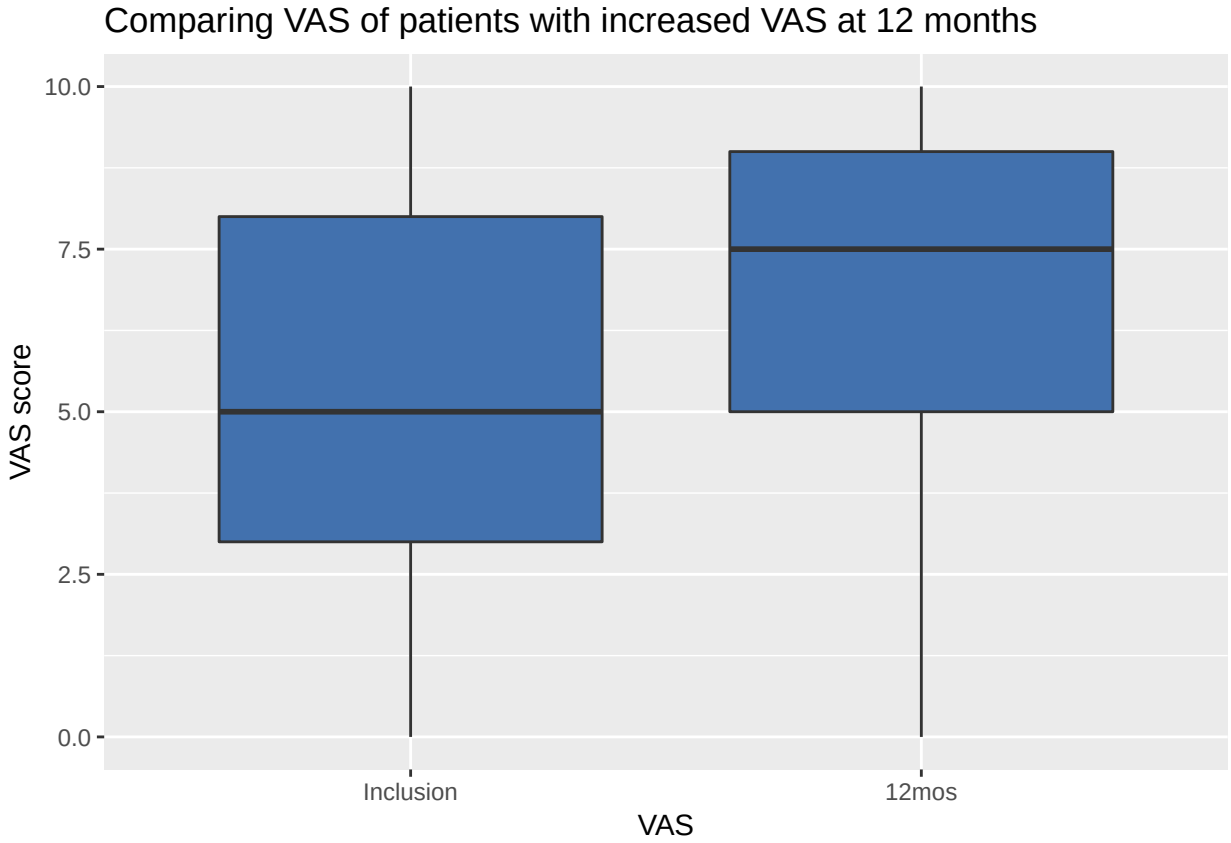
Reference for scaling

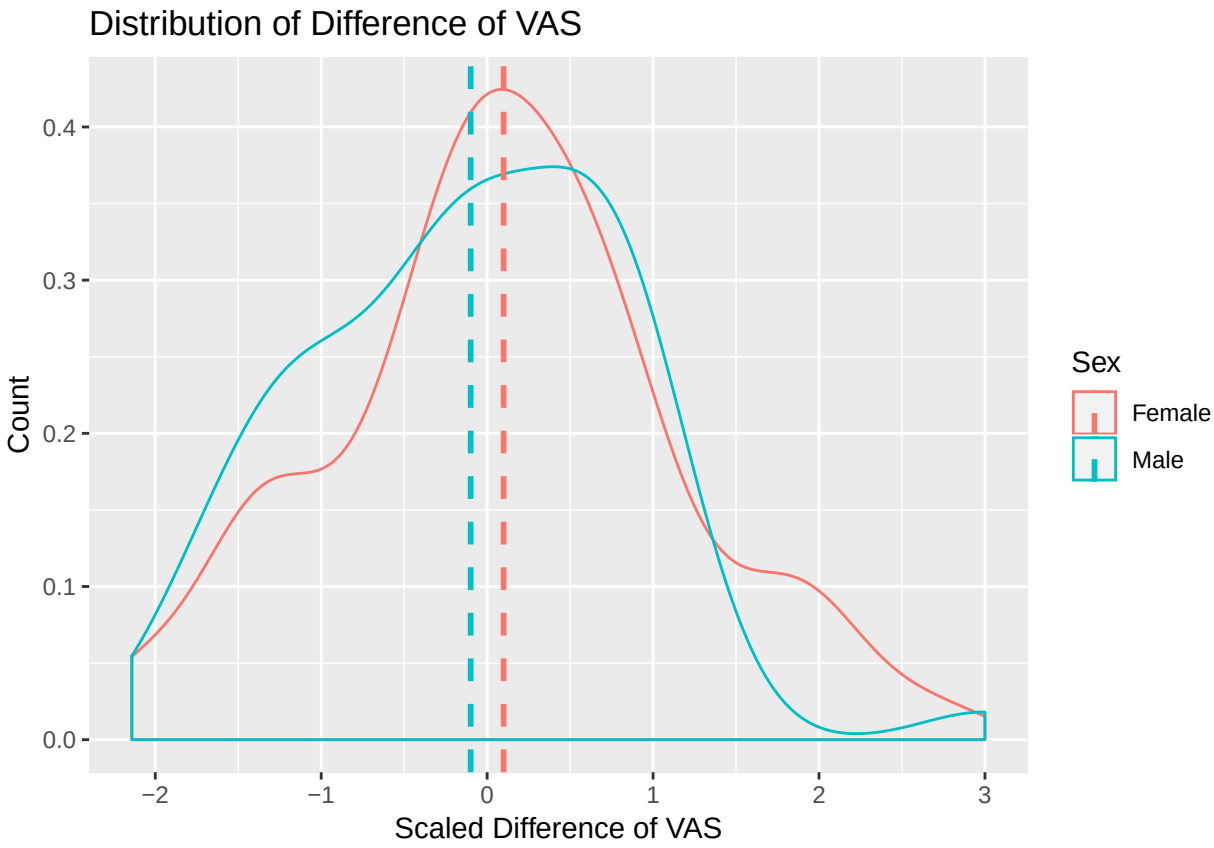
Table 6: Patients with increased VAS after 12 months

patientID	Age	Sex	Smoker	Inclusion	12mos	PainDiff
1	56	Male	Non-Smoker	3.0	4.0	1.0
5	25	Female	Non-Smoker	3.0	3.9	0.9
6	39	Male	Non-Smoker	3.5	5.0	1.5
14	41	Male	Non-Smoker	2.0	4.0	2.0
23	52	Female	Smoker	8.0	8.0	0.0
32	32	Male	Smoker	8.0	9.5	1.5
45	28	Female	Non-Smoker	8.0	9.1	1.1
55	49	Female	Smoker	5.0	9.0	4.0
60	41	Male	Non-Smoker	9.0	9.0	0.0
71	59	Female	Non-Smoker	3.0	8.0	5.0
76	57	Female	Non-Smoker	4.0	5.0	1.0
78	45	Female	Non-Smoker	5.0	5.0	0.0
81	30	Female	Smoker	2.5	9.5	7.0
83	37	Female	Non-Smoker	3.5	7.0	3.5
84	38	Male	Smoker	7.5	8.3	0.8
85	31	Male	Non-Smoker	10.0	10.0	0.0
86	32	Female	Non-Smoker	2.0	6.5	4.5
90	49	Female	Smoker	6.0	7.5	1.5
92	33	Male	Smoker	8.0	8.3	0.3
94	38	Male	Non-Smoker	5.5	6.1	0.6

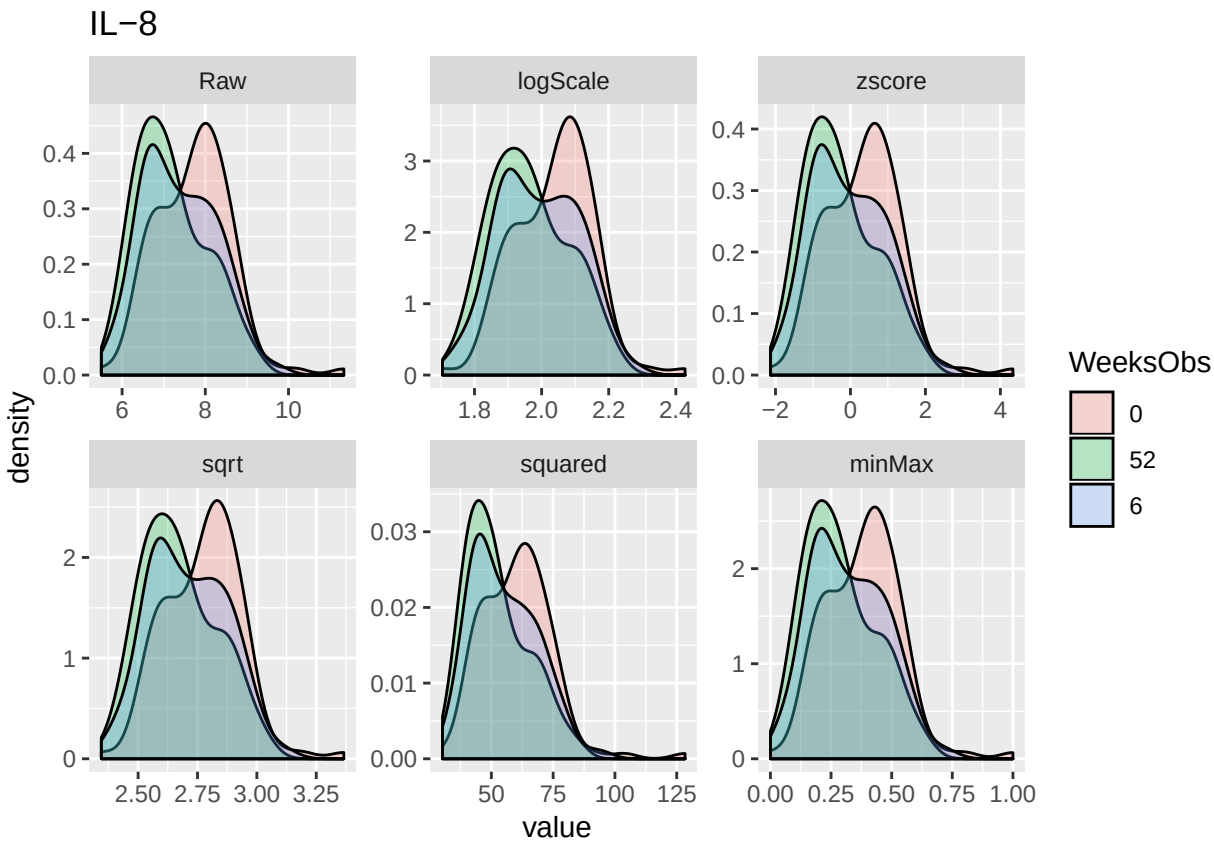
patientID	Age	Sex	Smoker	Inclusion	12mos	PainDiff
95	36	Female	Non-Smoker	3.0	3.0	0.0
99	59	Male	Non-Smoker	6.5	7.0	0.5
105	45	Male	Smoker	5.0	5.0	0.0
112	29	Female	Non-Smoker	8.0	9.0	1.0
114	50	Male	Non-Smoker	0.0	0.0	0.0
118	52	Female	Smoker	3.0	3.0	0.0
127	35	Male	Non-Smoker	9.0	10.0	1.0
135	43	Female	Smoker	2.0	4.5	2.5
140	52	Male	Smoker	8.5	9.0	0.5
149	46	Female	Non-Smoker	5.0	9.0	4.0
150	57	Male	Smoker	1.0	9.0	8.0

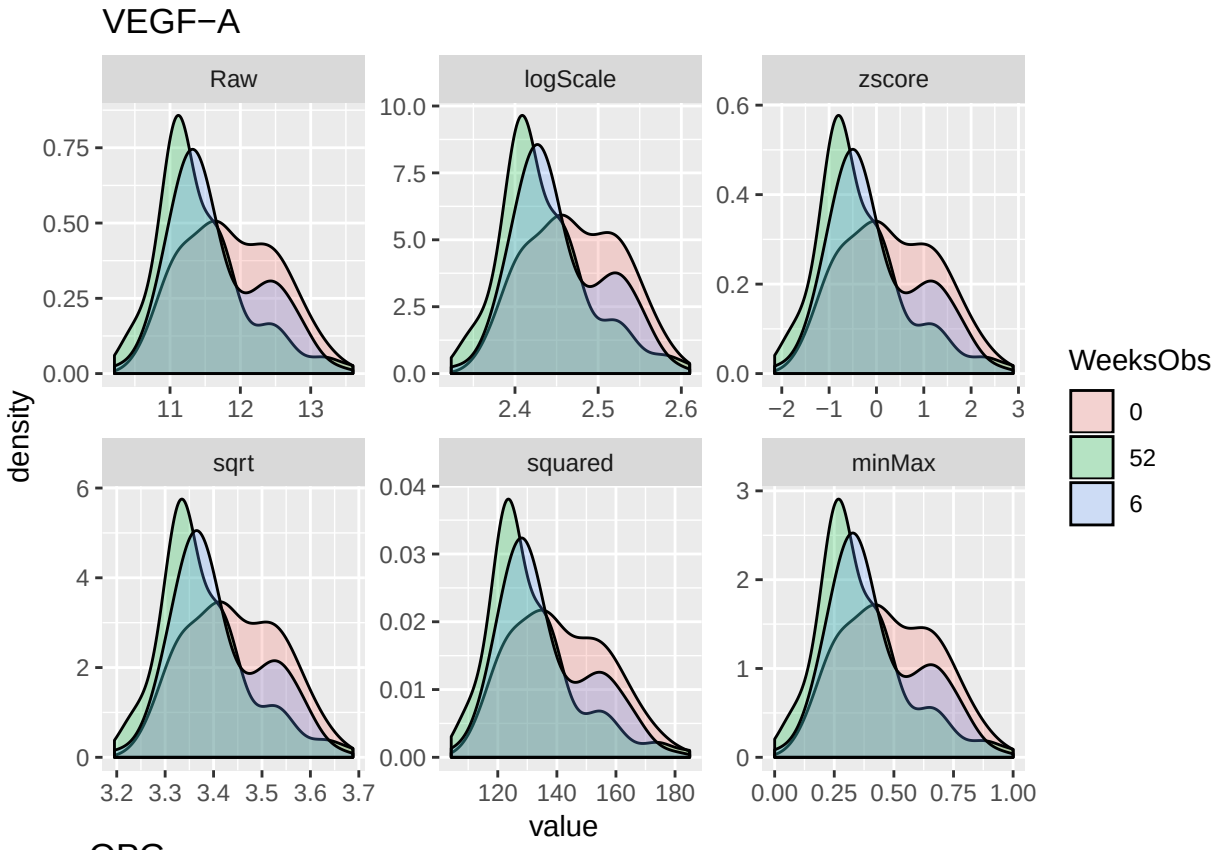
31 patients reported to have higher VAS compared to inclusion. 16 female patients and 15 male patients infomed drugs were not helpful 12 Smokers and
19 non smokers

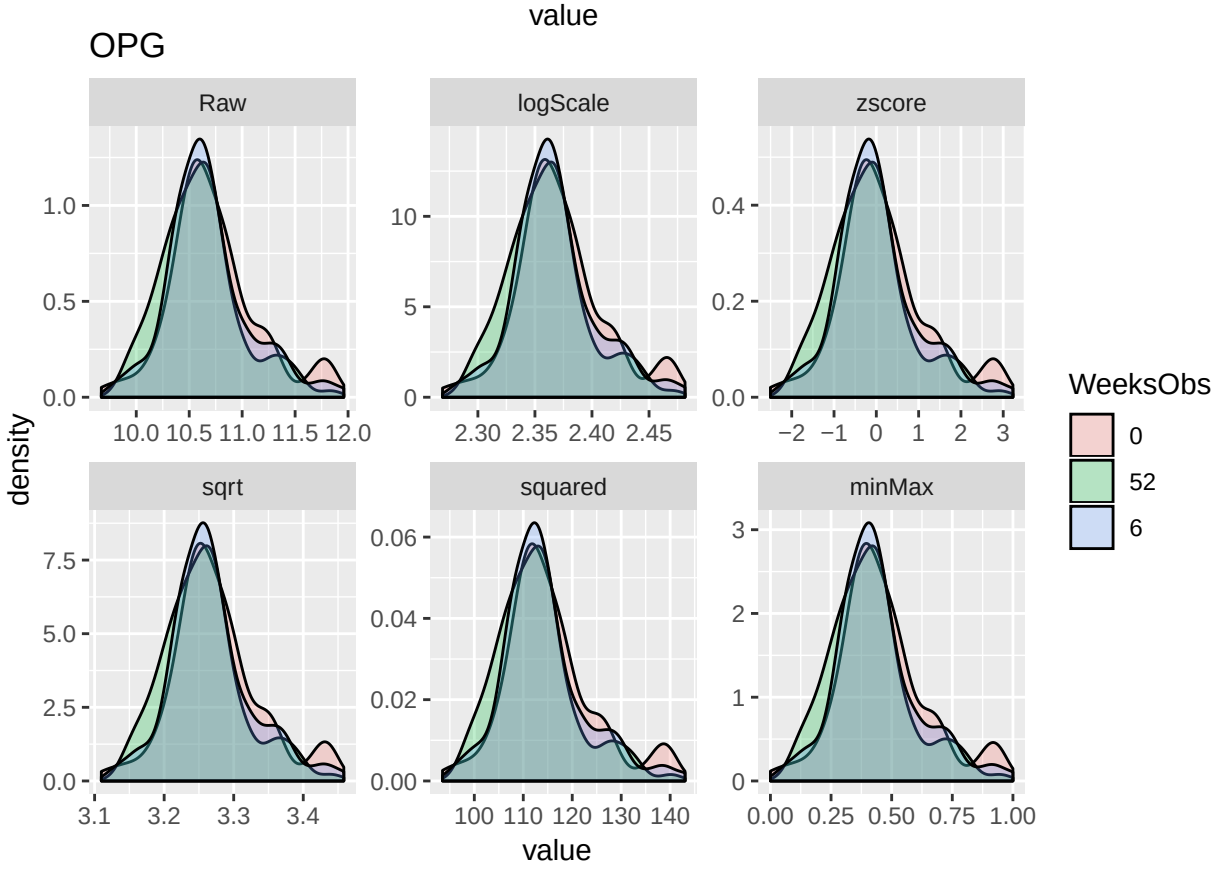


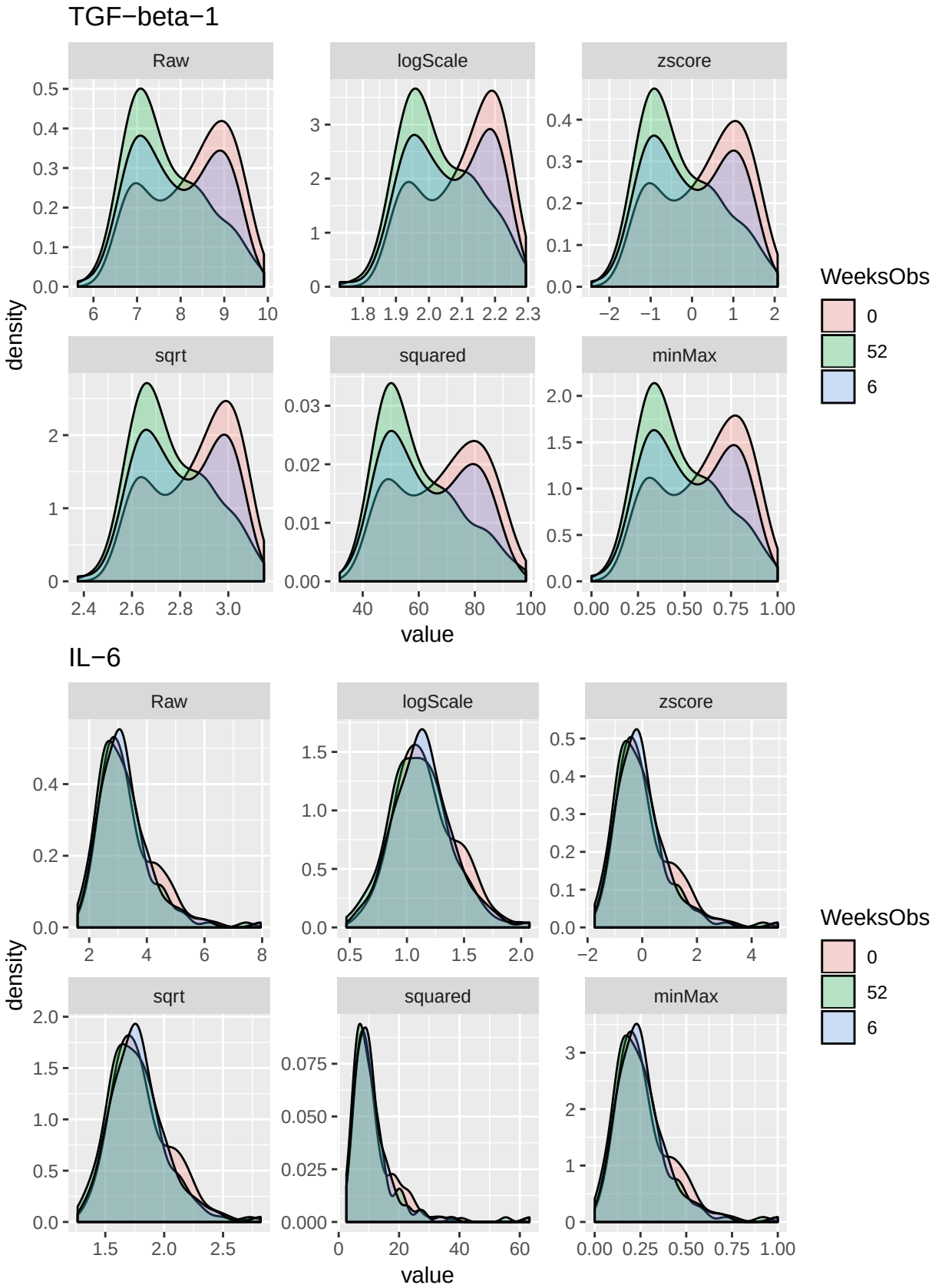

Distribution of Difference of VAS
Count
0.4
0.3
0.2
0.1
0.0
−2
−1
0
1
2
3
Scaled Difference of VAS
Sex
Female
Male

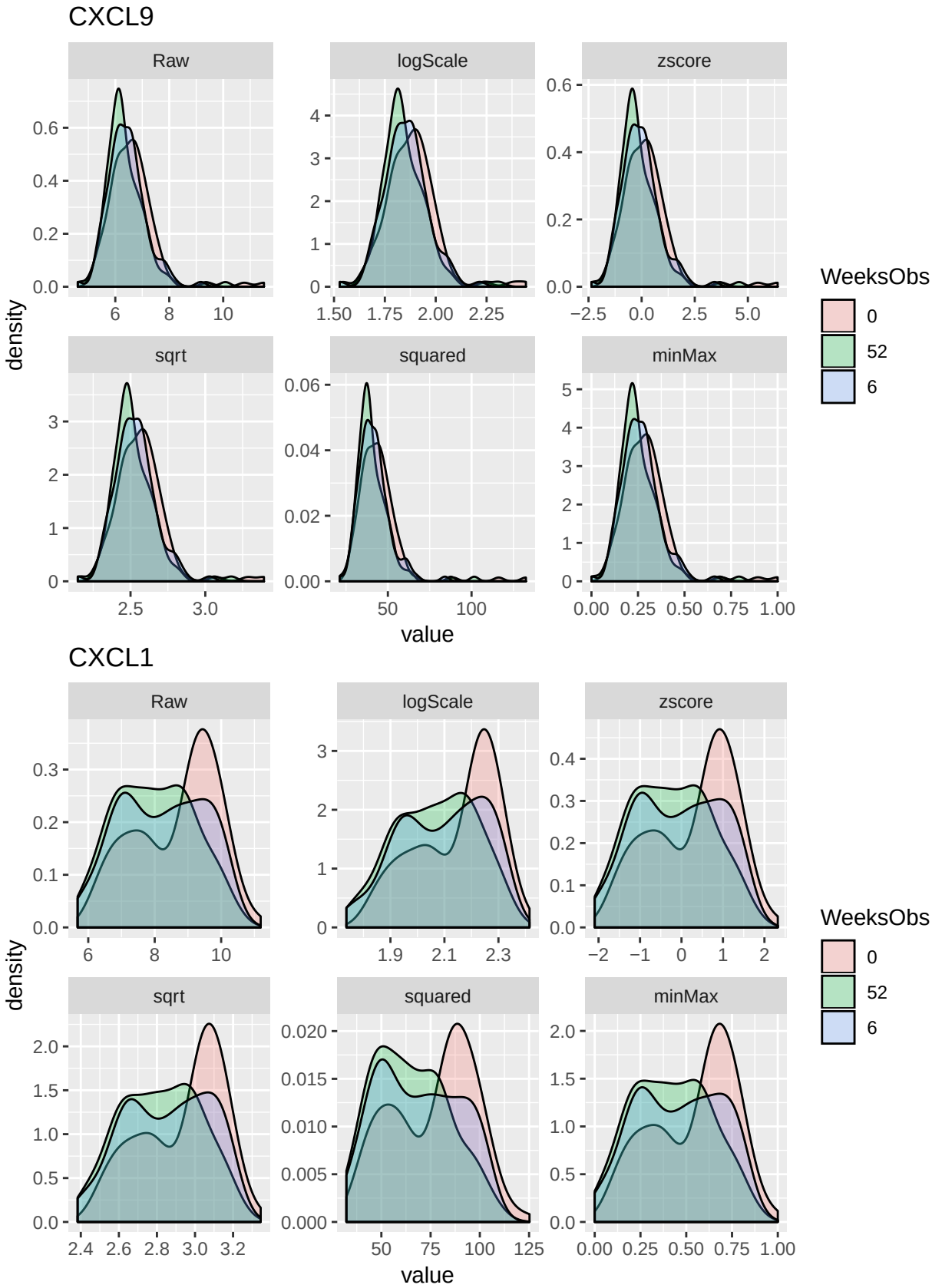
Transform Biomarkers

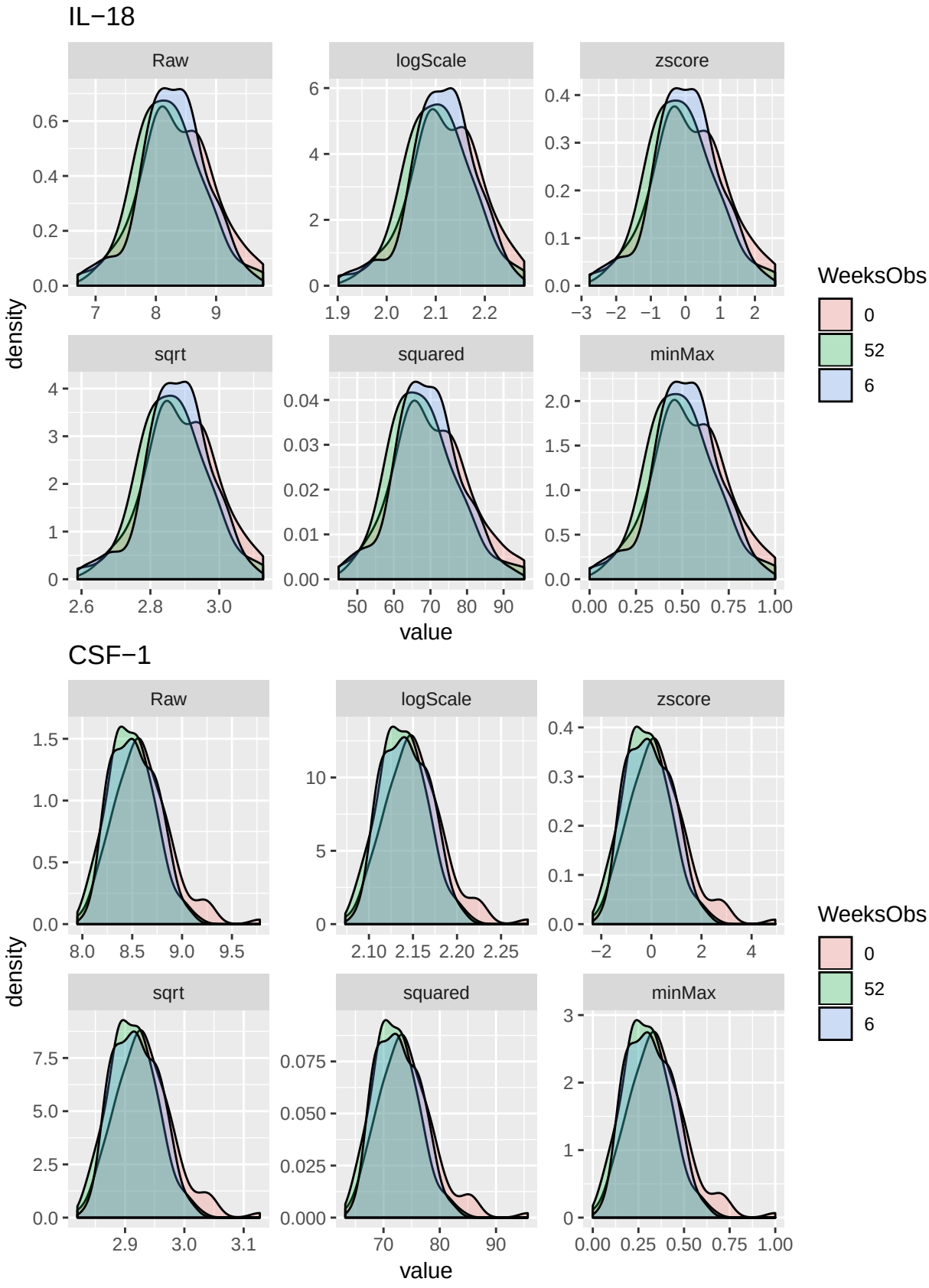



VEGF−A
Raw
logScale
zscore
sqrt
squared
minMax
density
value
WeeksObs
0
52
6


OPG
Raw
logScale
zscore
sqrt
squared
minMax
density
value
WeeksObs
0
52
6


TGF−beta−1
Raw
logScale
zscore
sqrt
squared
minMax
density
value
WeeksObs
0
52
6
IL−6
Raw
logScale
zscore
sqrt
squared
minMax
density
value
WeeksObs
0
52
6


CXCL9
Raw
logScale
zscore
sqrt
squared
minMax
density
value
WeeksObs
0
52
6
CXCL1
Raw
logScale
zscore
sqrt
squared
minMax
density
value
WeeksObs
0
52
6



Based on all of these graphs, zscore transformation seems to be the most suitable transformation which

provides us Normal distribution.