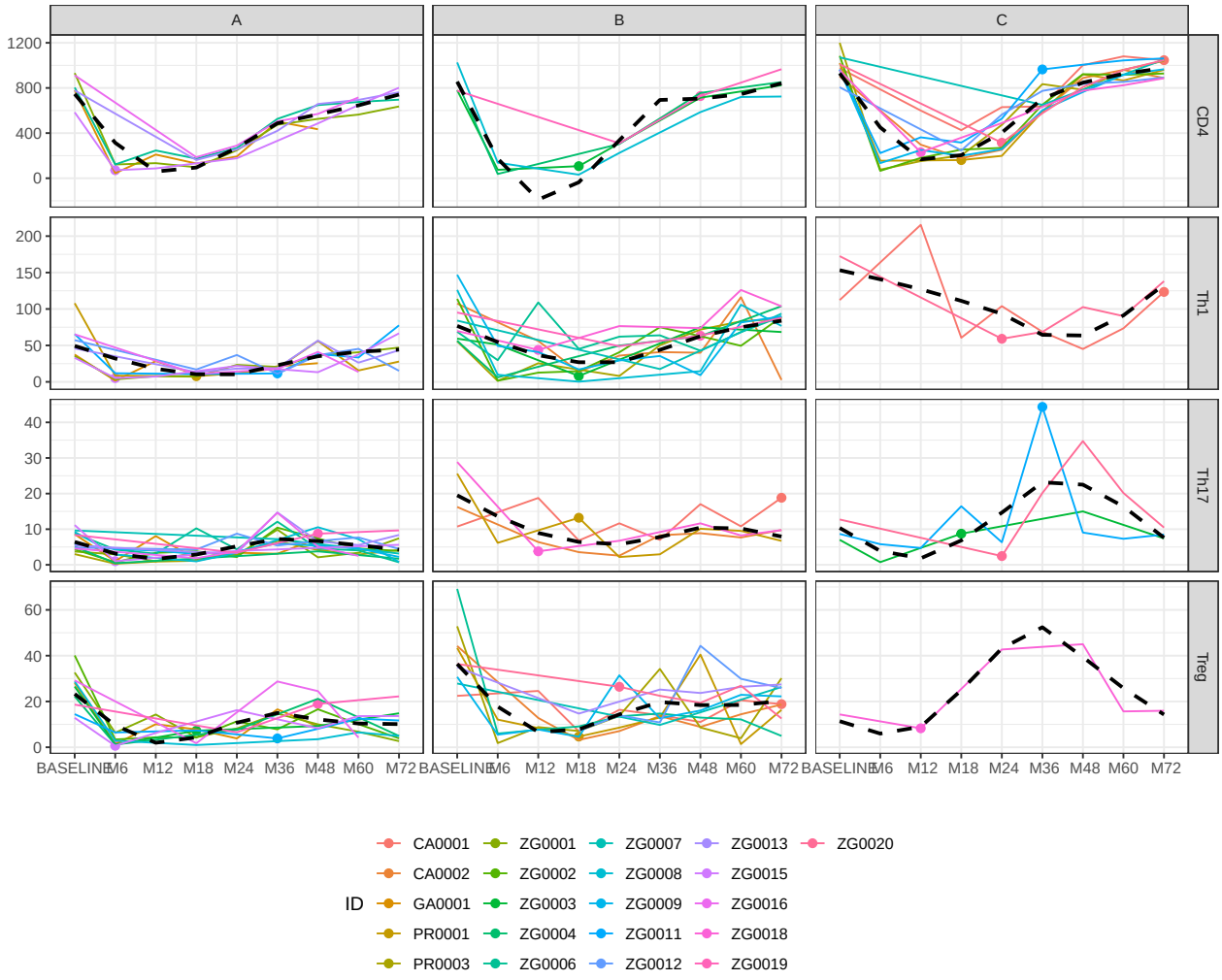
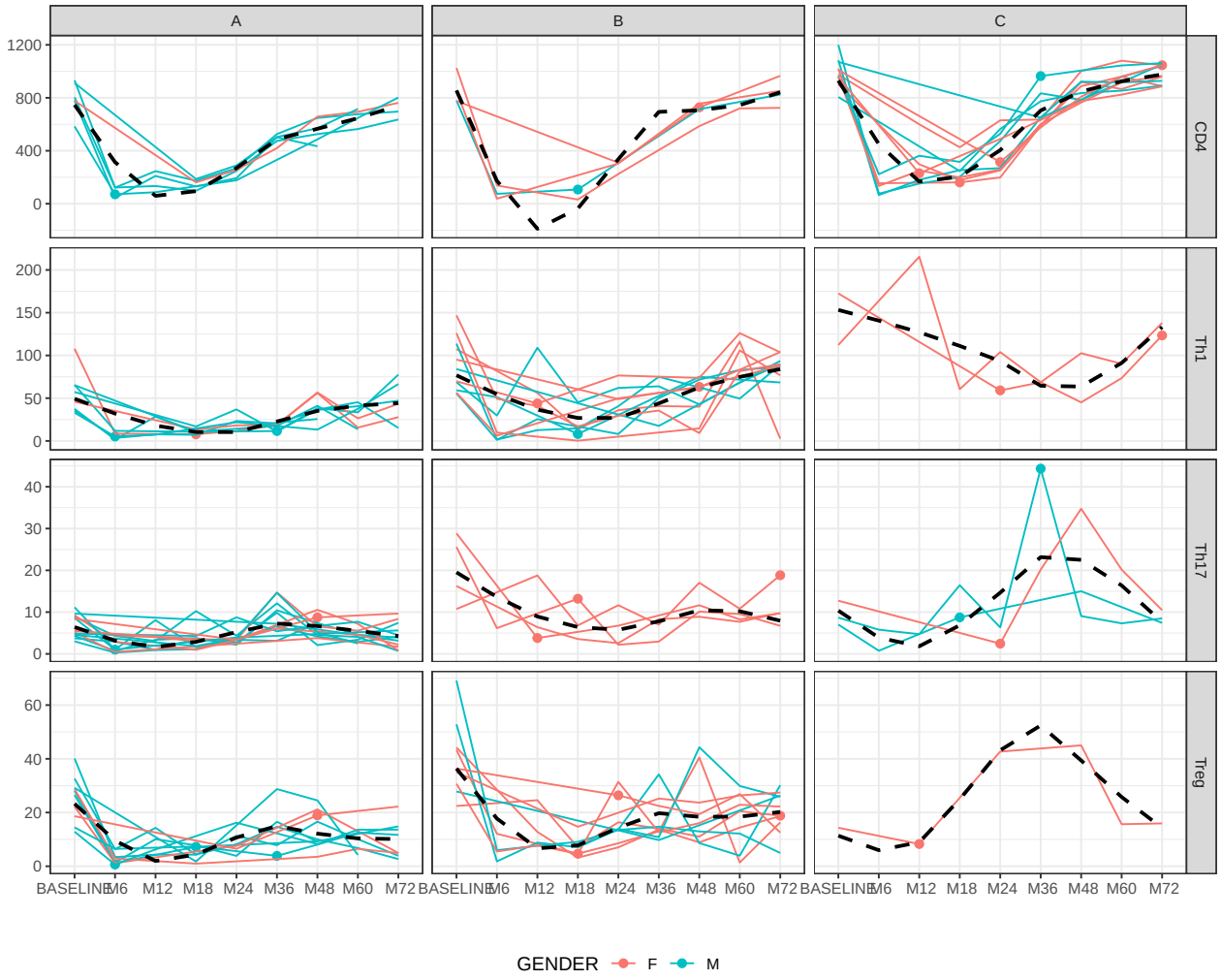
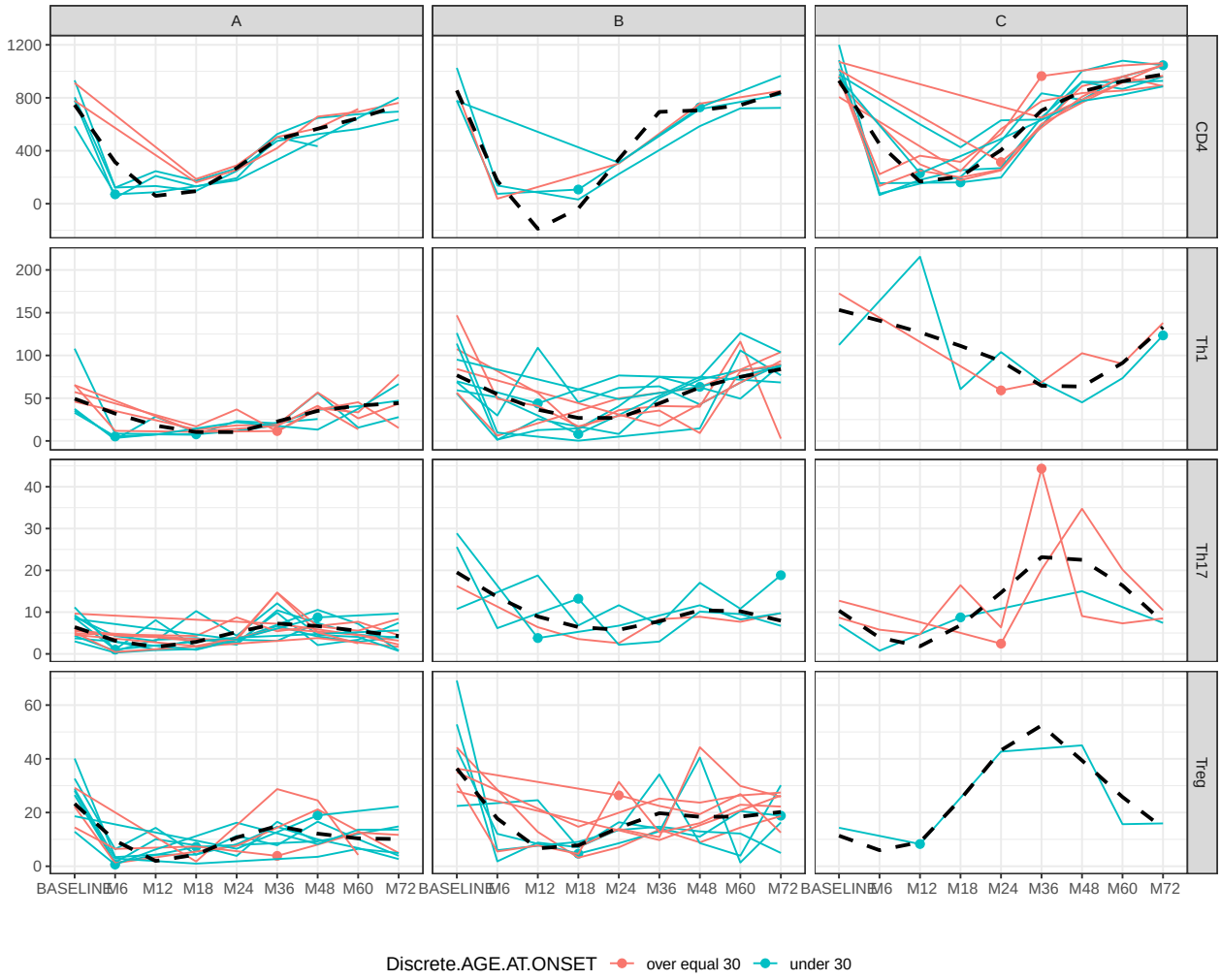
Data Investigation

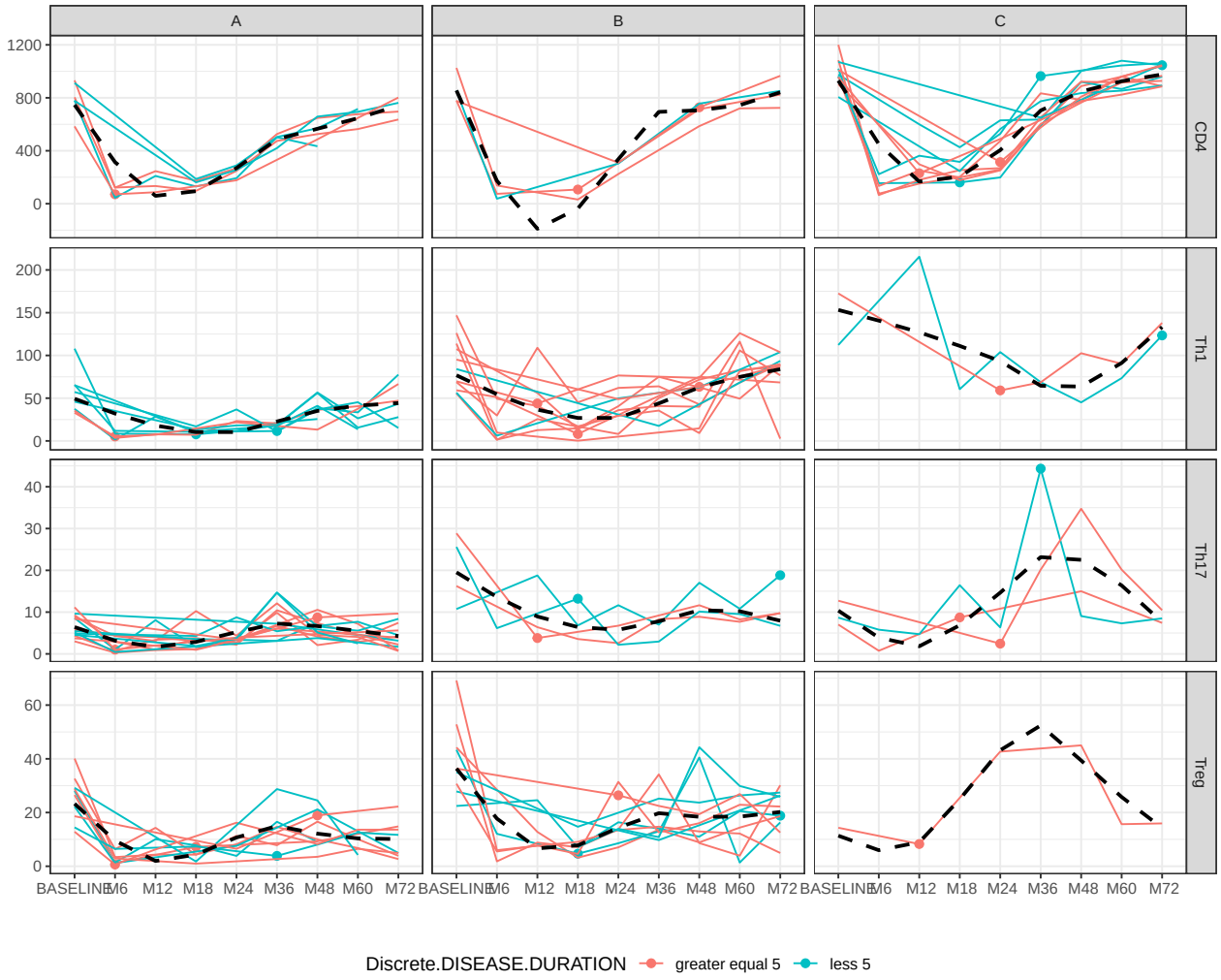
Connector Clustering

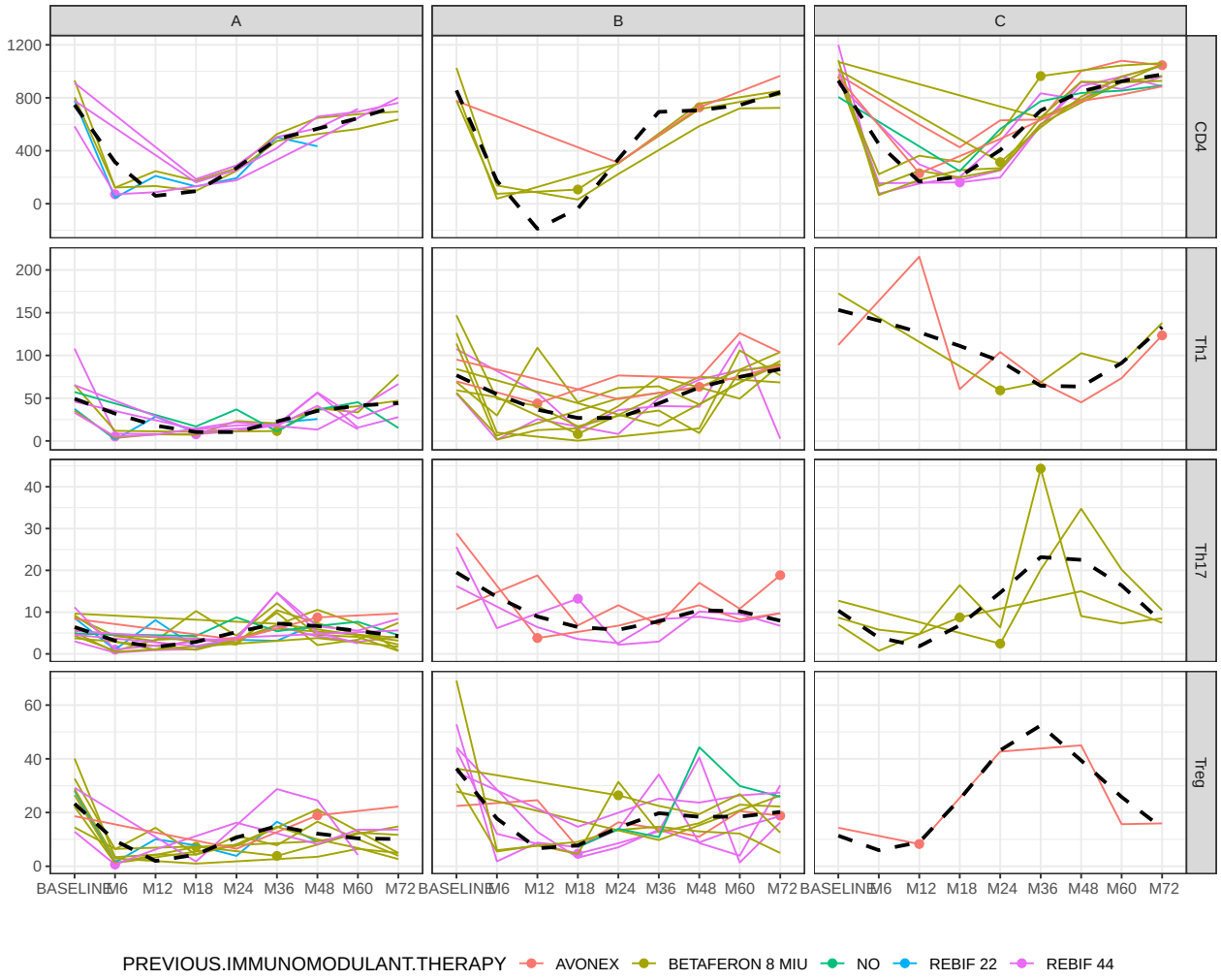
I primi plot rappresentano la divisione in cluster ottenuti usando CONNECTOR (le colonne) per ogni esperimento (le righe). I colori dipendono dai pazienti, il numero di relapse (in ordine).

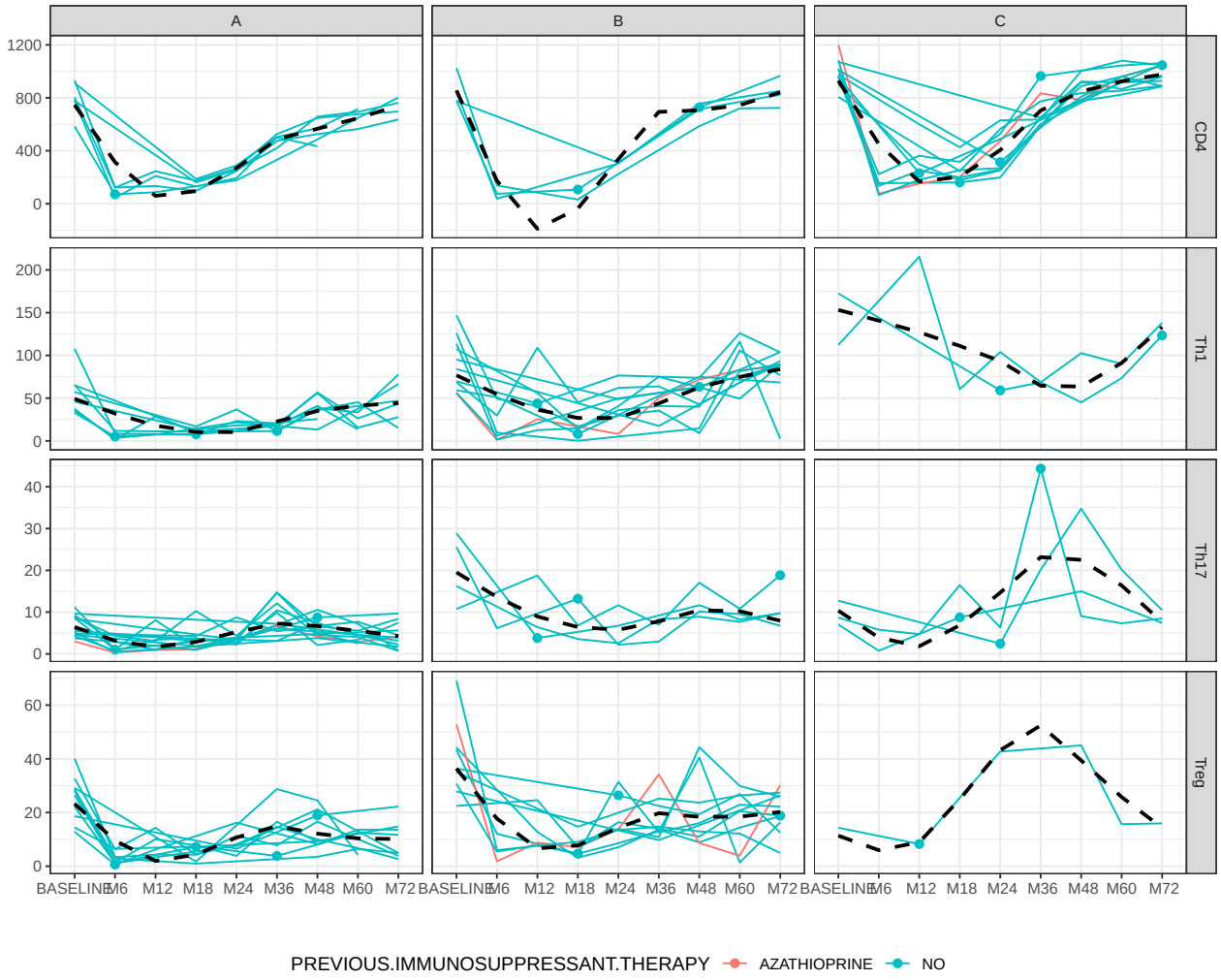


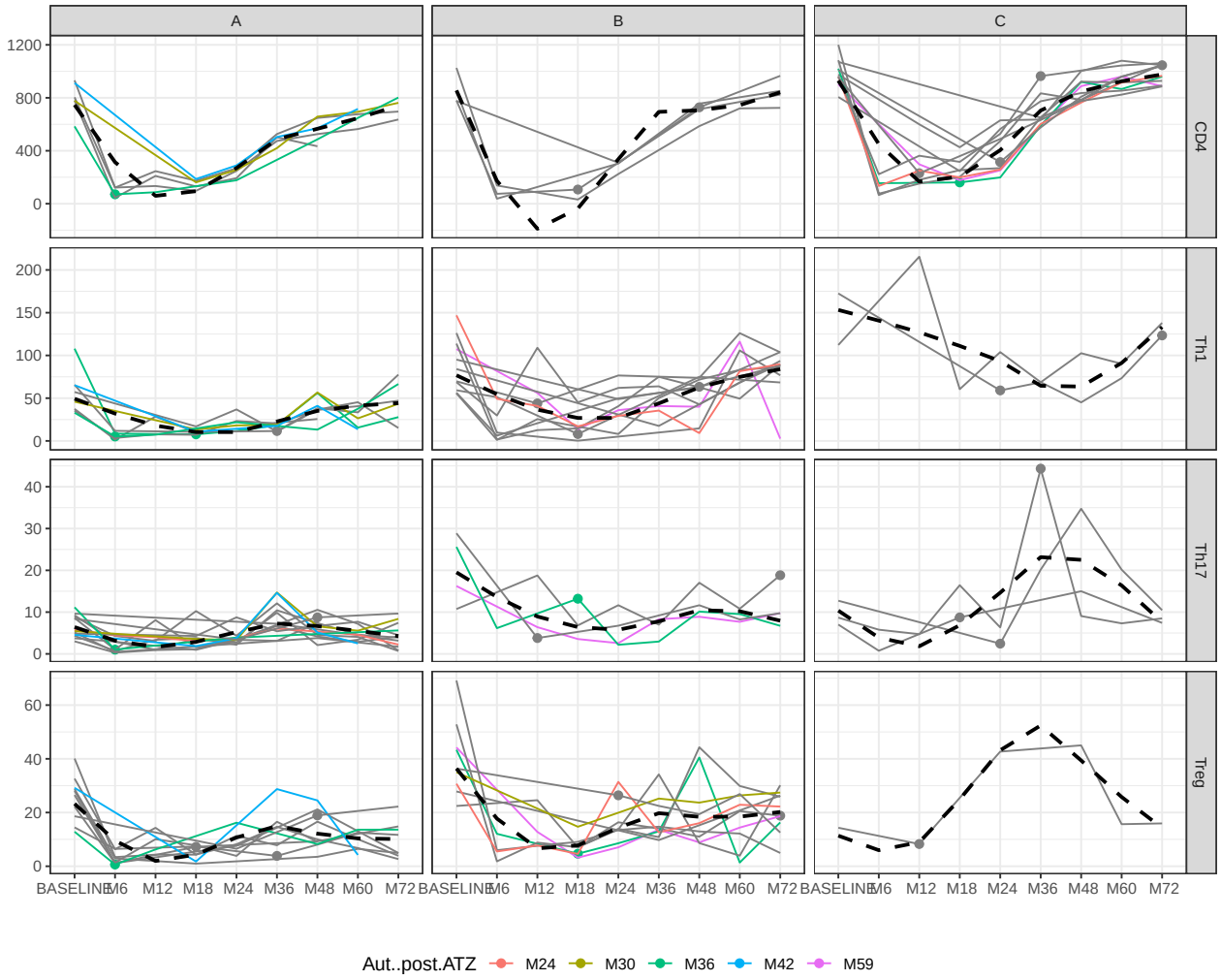

A
B
C
CD4
Th1
Th17
Treg
1200
800
400
0
200
150
100
50
40
30
20
10
60
BASELINE
M6
M12
M18
M24
M36
M48
M60
M72
GENDER
F
M

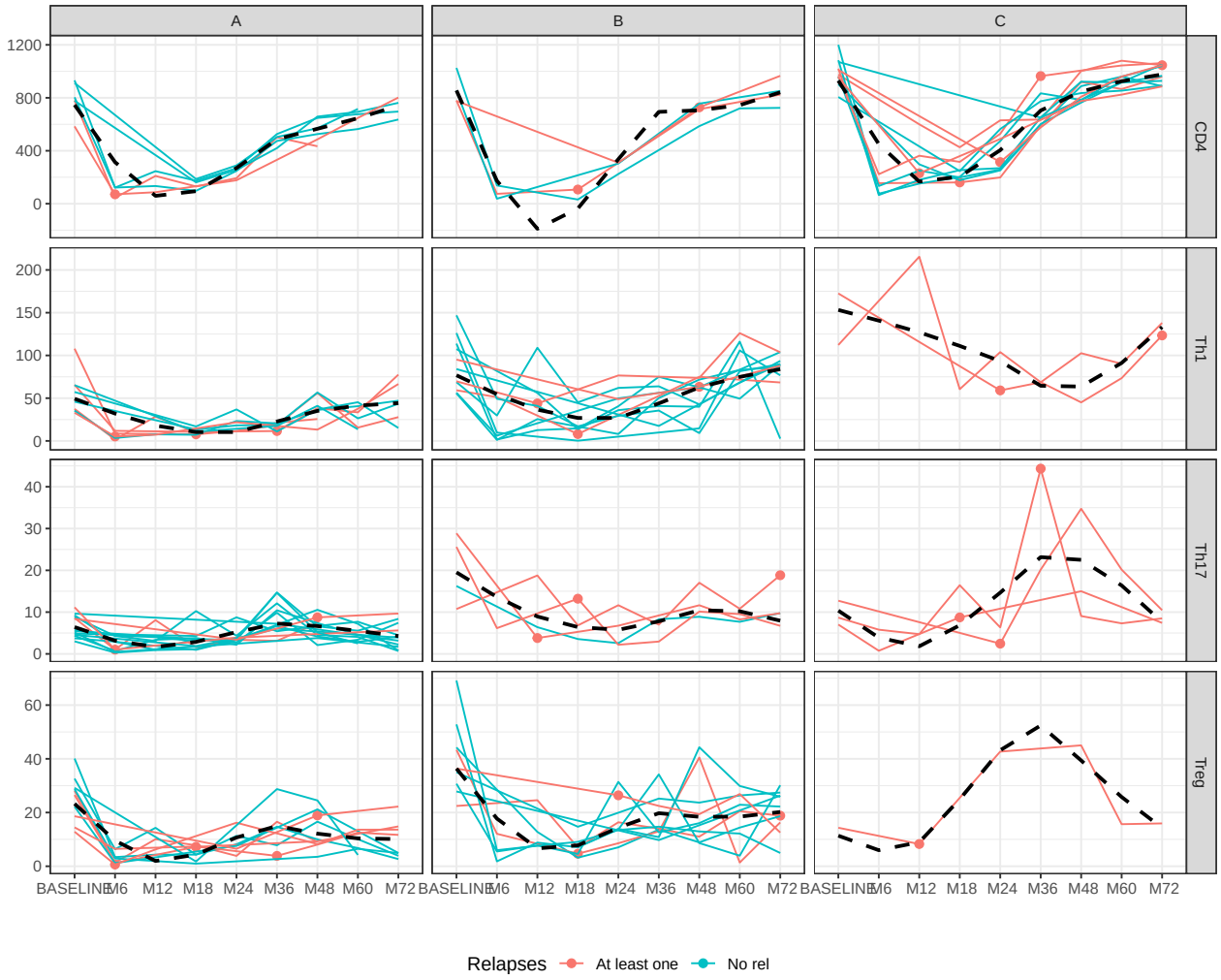

A
B
C
CD4
Th1
Th17
Treg
BASELINE
M6
M12
M18
M24
M36
M48
M60
M72
Discrete.AGE.AT.ONSET
over equal 30
under 30

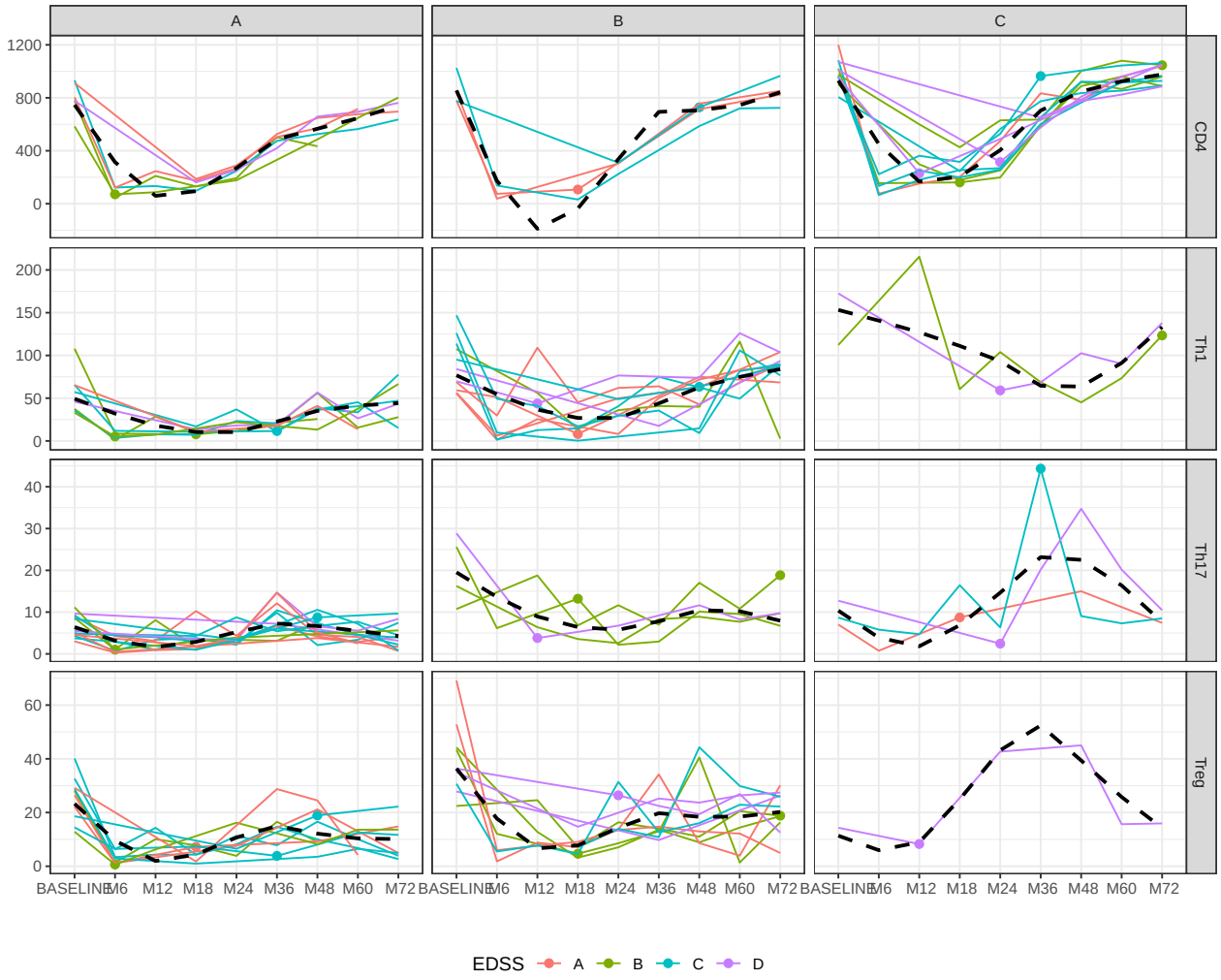

A
B
C
CD4
Th1
Th17
Treg
BASELINE
M6
M12
M18
M24
M36
M48
M60
M72
Discrete.DISEASE.DURATION
greater equal 5
less 5

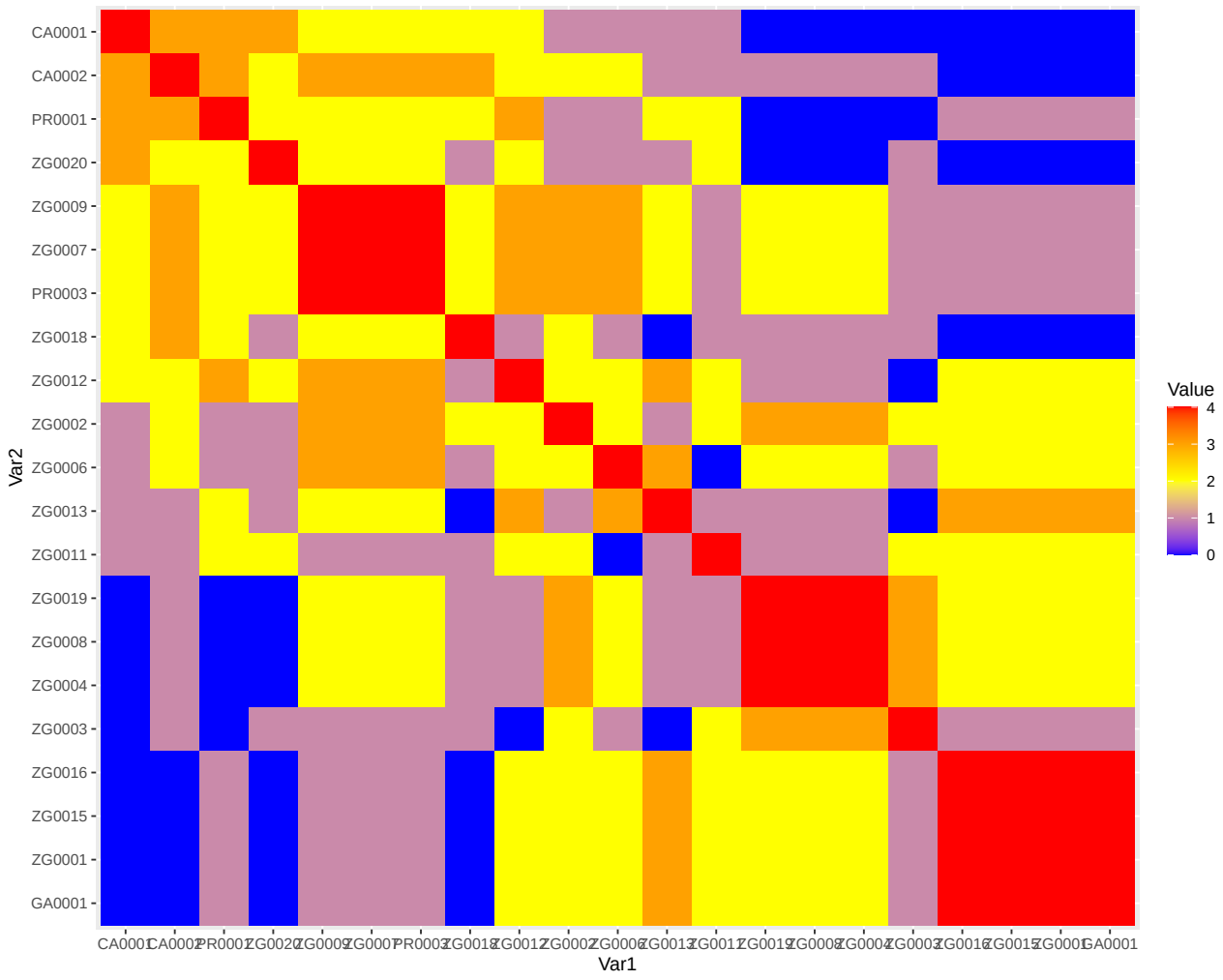

A
B
C
CD4
Th1
Th17
Treg
BASELINE
M6
M12
M18
M24
M36
M48
M60
M72
PREVIOUS.IMMUNOMODULANT.THERAPY
AVONEX
BETAFERON 8 MIU
NO
REBIF 22
REBIF 44


A
B
C
CD4
Th1
Th17
Treg
BASELINE
M6
M12
M18
M24
M36
M48
M60
M72
PREVIOUS.IMMUNOSUPPRESSANT.THERAPY
AZATHIOPRINE
NO


A
B
C
CD4
Th1
Th17
Treg
BASELINE
M6
M12
M18
M24
M36
M48
M60
M72
Aut..post.ATZ
M24
M30
M36
M42
M59


A
B
C
CD4
Th1
Th17
Treg
1200
800
400
0
200
150
100
50
40
30
20
10
60
BASELINE
M6
M12
M18
M24
M36
M48
M60
M72
Relapses
At least one
No rel


A
B
C
CD4
Th1
Th17
Treg
BASELINE
M6
M12
M18
M24
M36
M48
M60
M72
EDSS
A
B
C
D



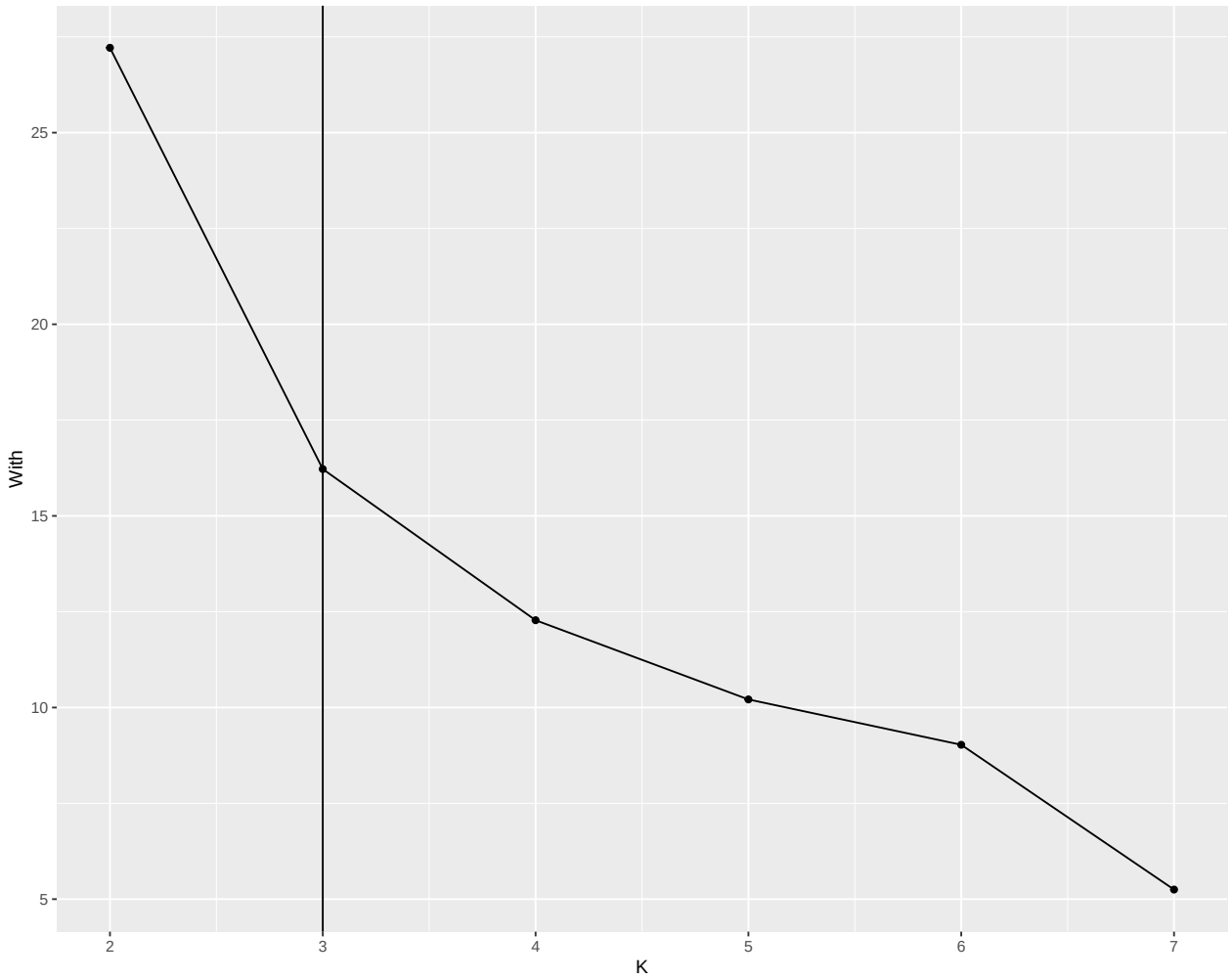
kmeans Clustering

Provo ad unire tutti i cluster che ho ottenuto nei vari esperimenti e ne faccio un secondo clustering (usendo il metodo di kmeans algorithm) per trovare i gruppi di pazienti (meta-pazienti) che hanno clusters di connector simili.

```
## Loading required package: MASS
```

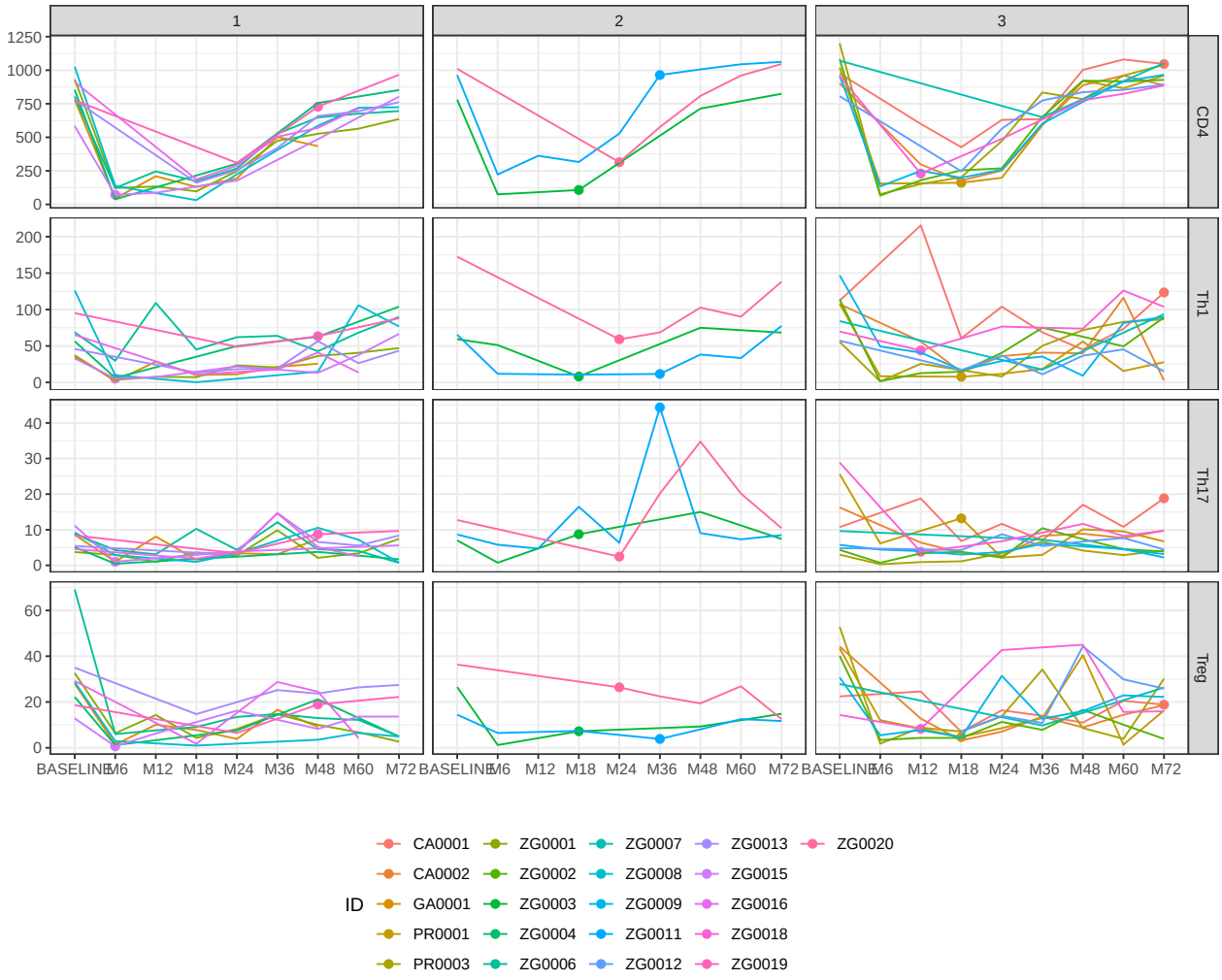
```
##
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:dplyr':
##
##     select
```



ID	CD4	Th1	Th17	Treg	ClusterKmeans
CA0001	C	C	B	B	3
CA0002	C	B	B	B	3
GA0001	A	A	A	A	1
PR0001	C	A	B	B	3
PR0003	C	B	A	B	3
ZG0001	A	A	A	A	1
ZG0002	C	B	A	A	3
ZG0003	B	B	C	A	2
ZG0004	B	B	A	A	1
ZG0006	A	B	A	B	1
ZG0007	C	B	A	B	3
ZG0008	B	B	A	A	1
ZG0009	C	B	A	B	3
ZG0011	C	A	C	A	2
ZG0012	C	A	A	B	3
ZG0013	A	A	A	B	1
ZG0015	A	A	A	A	1
ZG0016	A	A	A	A	1
ZG0018	C	B	B	C	3
ZG0019	B	B	A	A	1

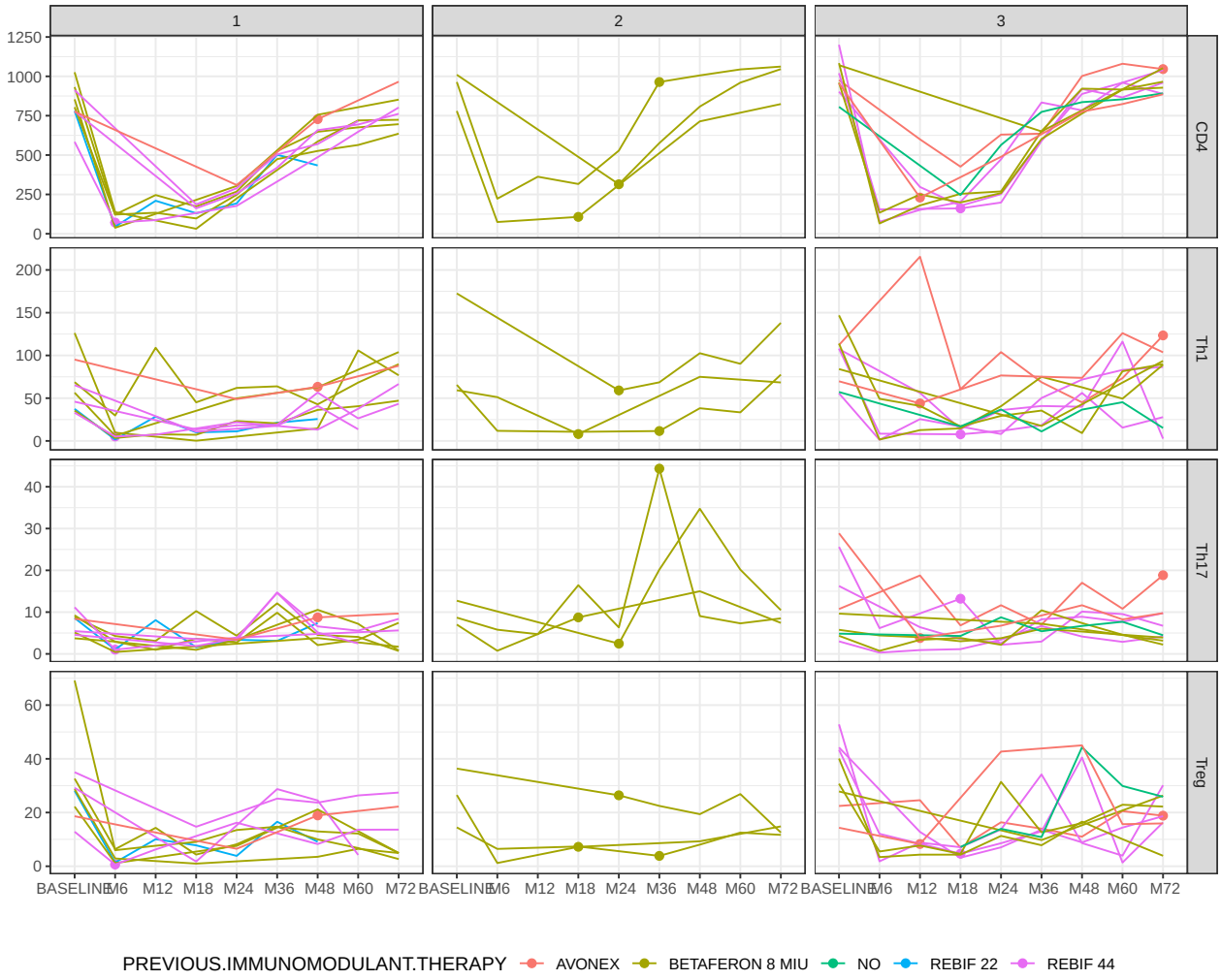
ID	CD4	Th1	Th17	Treg	ClusterKmeans
ZG0020	C	C	C	B	2

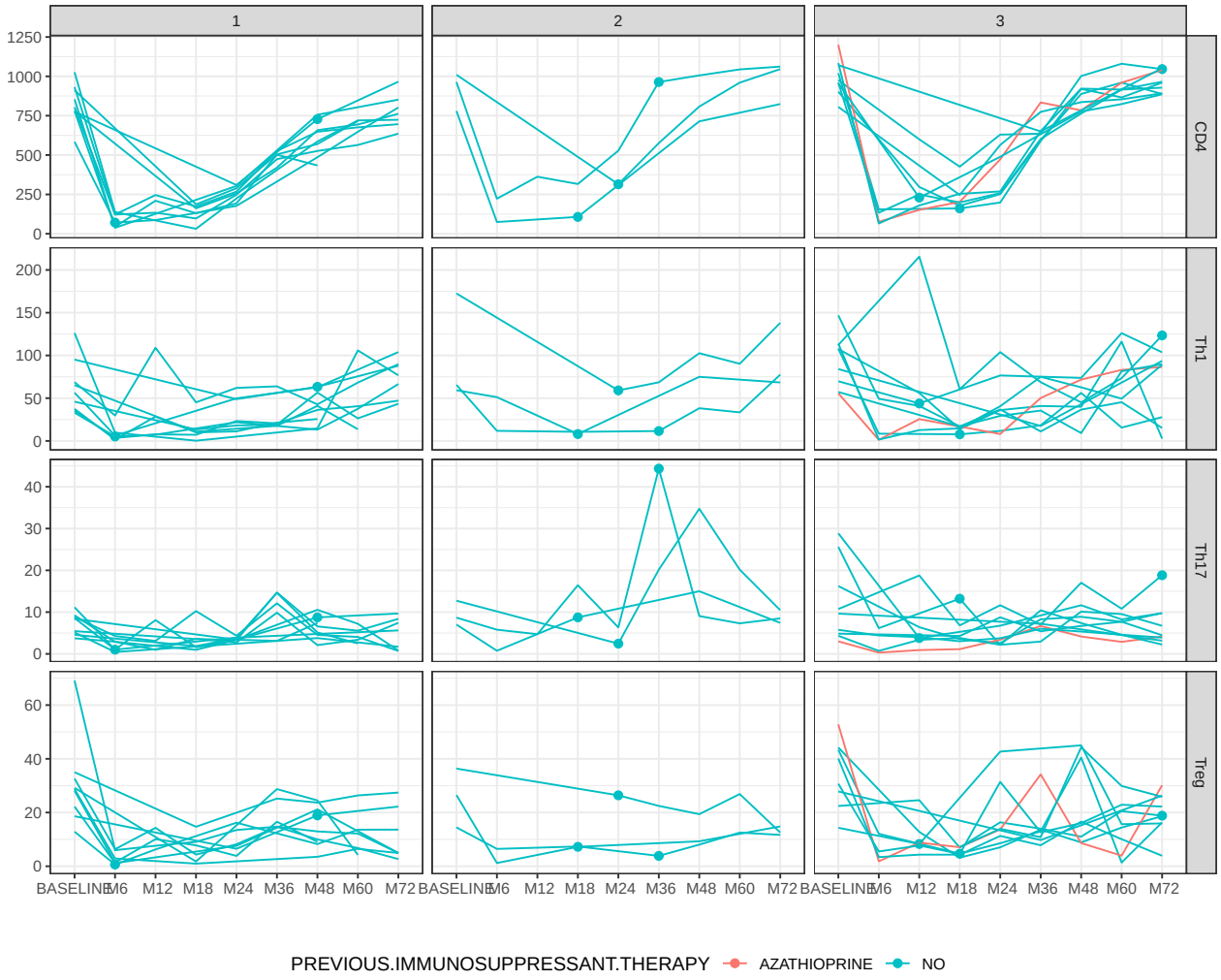


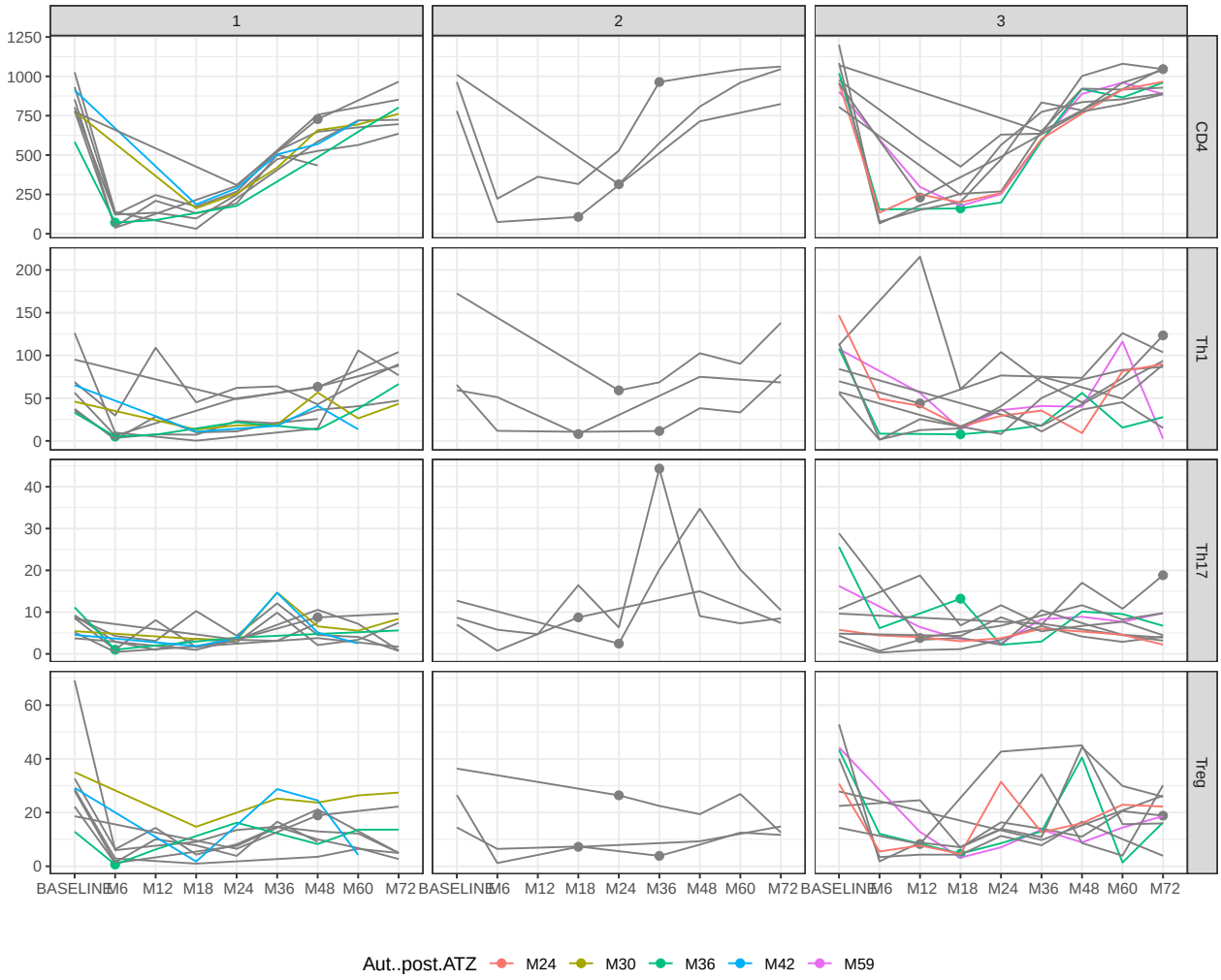
1
2
3
CD4
Th1
Th17
Treg
1250
1000
750
500
250
0
200
150
100
50
40
30
20
10
60
BASELINE
M6
M12
M18
M24
M36
M48
M60
M72
GENDER
F
M

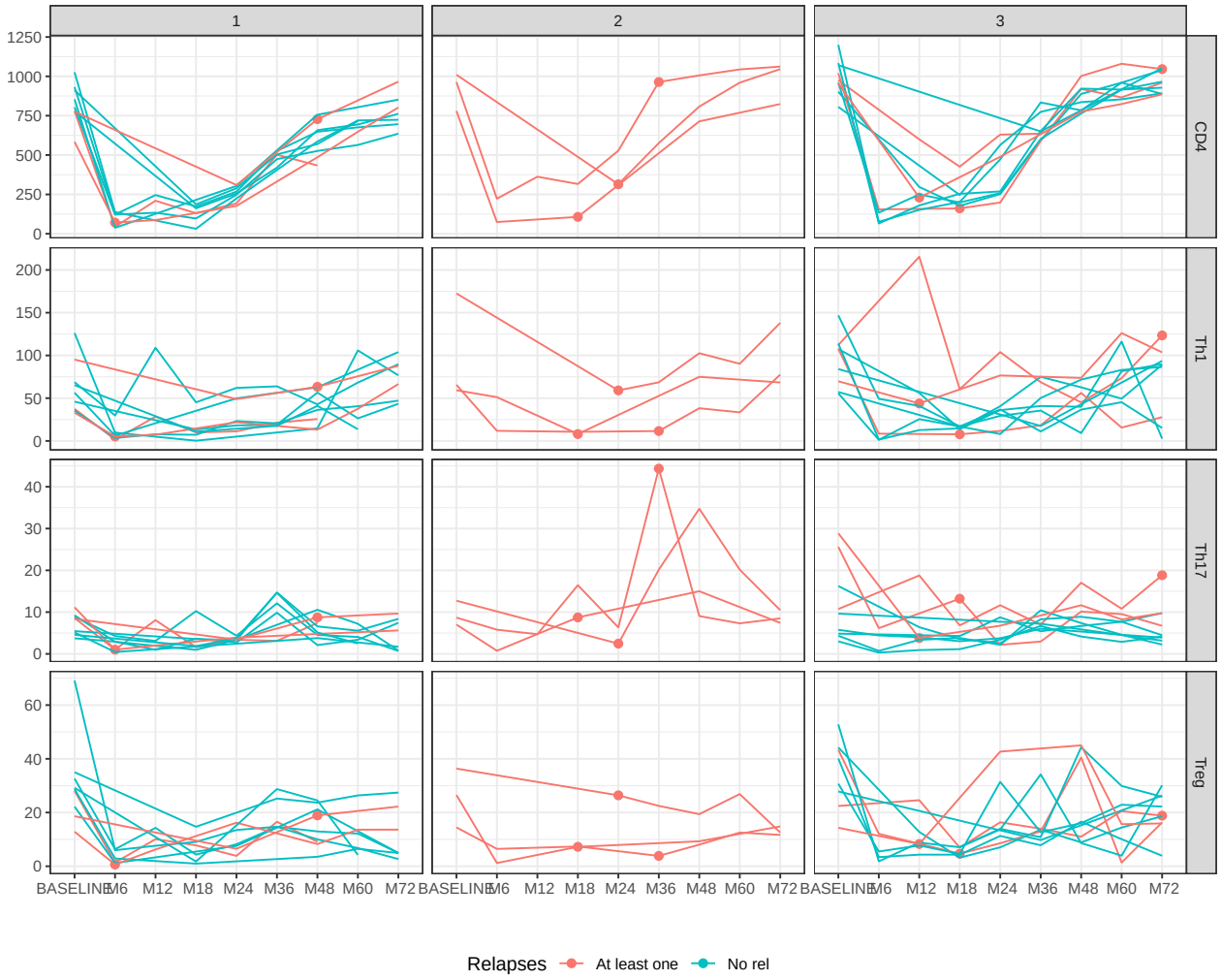
1
2
3
CD4
Th1
Th17
Treg
BASELINE
M6
M12
M18
M24
M36
M48
M60
M72
Discrete.AGE.AT.ONSET
over equal 30
under 30

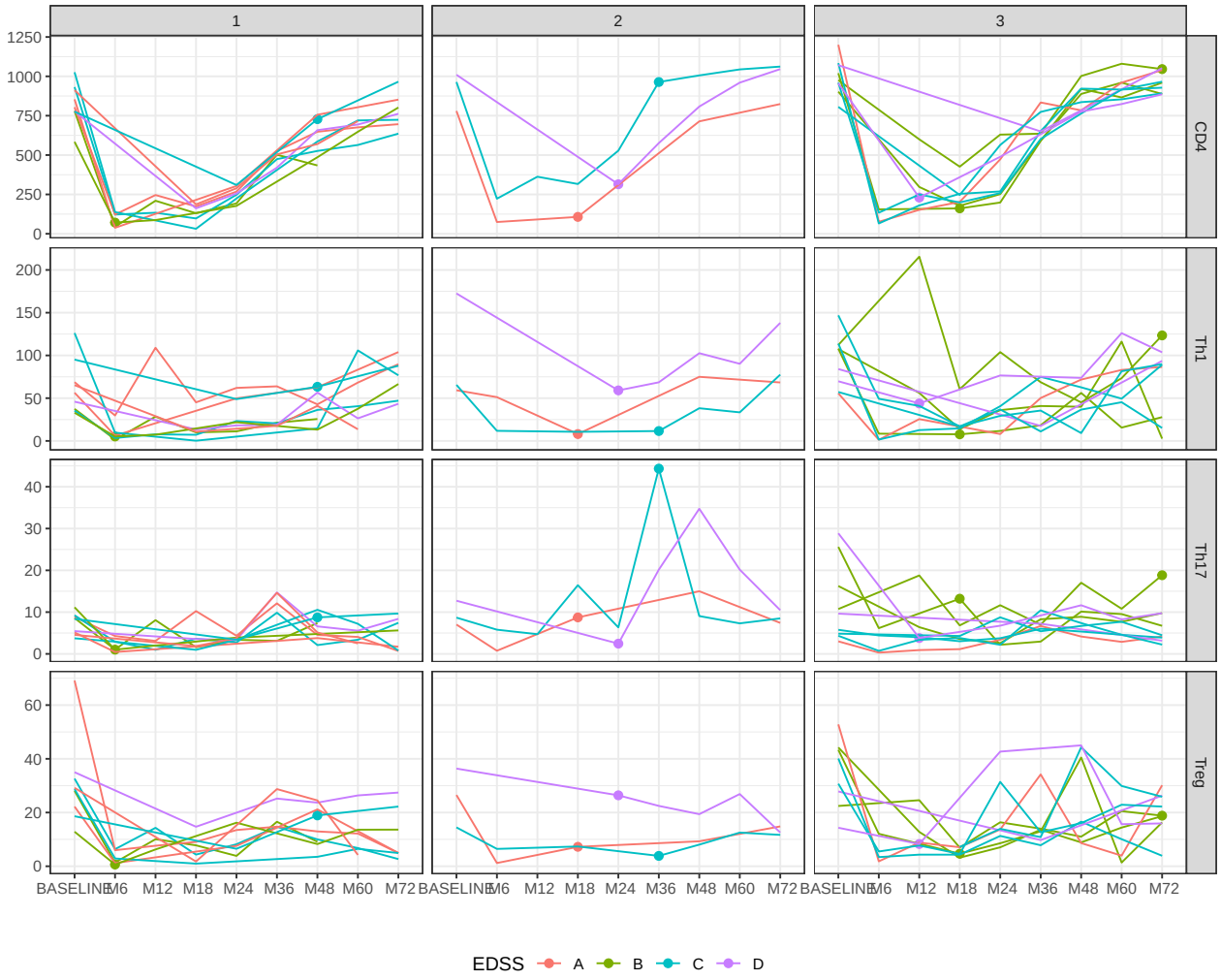
1
2
3
CD4
Th1
Th17
Treg
BASELINE
M6
M12
M18
M24
M36
M48
M60
M72
Discrete.DISEASE.DURATION
greater equal 5
less 5

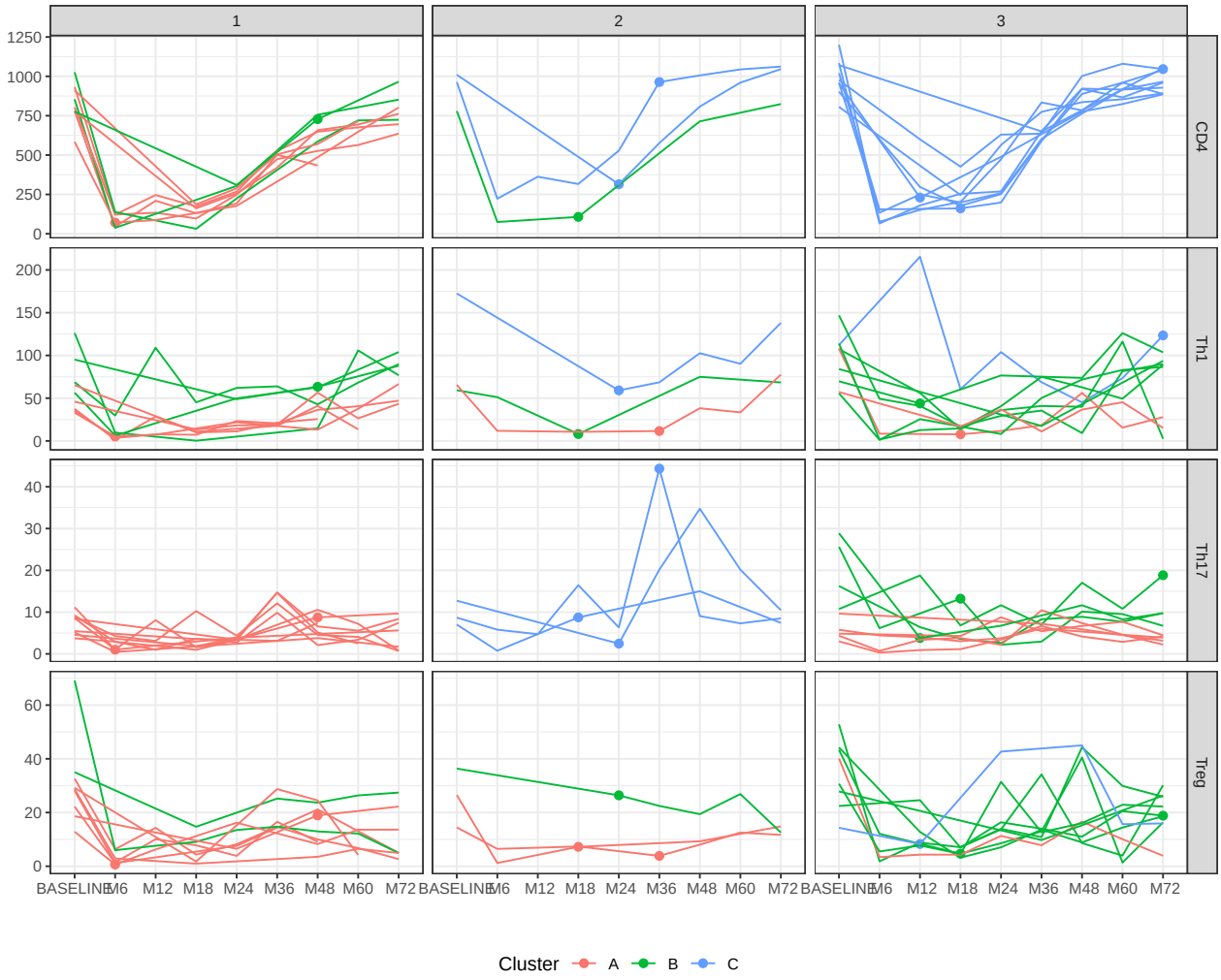

1
2
3
CD4
Th1
Th17
Treg
1250
1000
750
500
250
0
200
150
100
50
40
30
20
10
60
BASELINE
M6
M12
M18
M24
M36
M48
M60
M72
PREVIOUS.IMMUNOMODULANT.THERAPY
AVONEX
BETAFERON 8 MIU
NO
REBIF 22
REBIF 44


1
2
3
CD4
Th1
Th17
Treg
1250
1000
750
500
250
0
200
150
100
50
40
30
20
10
60
BASELINE
M6
M12
M18
M24
M36
M48
M60
M72
PREVIOUS.IMMUNOSUPPRESSANT.THERAPY
AZATHIOPRINE
NO


1
2
3
CD4
Th1
Th17
Treg
BASELINE
M6
M12
M18
M24
M36
M48
M60
M72
Aut..post.ATZ
M24
M30
M36
M42
M59

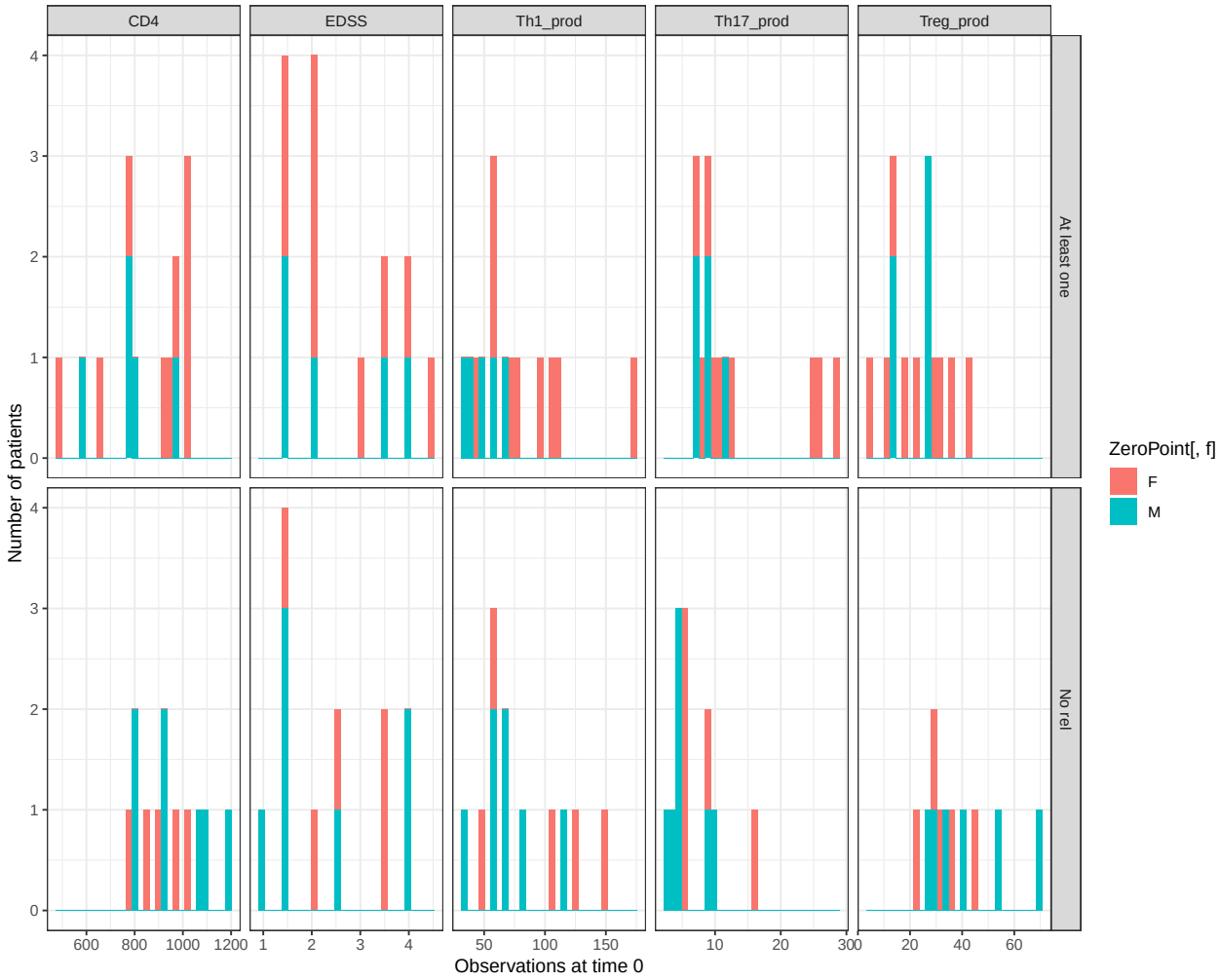

1
2
3
CD4
Th1
Th17
Treg
1250
1000
750
500
250
0
200
150
100
50
40
30
20
10
60
BASELINE
M6
M12
M18
M24
M36
M48
M60
M72
Relapses
At least one
No rel


1
2
3
CD4
Th1
Th17
Treg
BASELINE
M6
M12
M18
M24
M36
M48
M60
M72
EDSS
A
B
C
D


1
2
3
CD4
Th1
Th17
Treg
1250
1000
750
500
250
0
200
150
100
50
40
30
20
10
60
BASELINE
M6
M12
M18
M24
M36
M48
M60
M72
Cluster
A
B
C

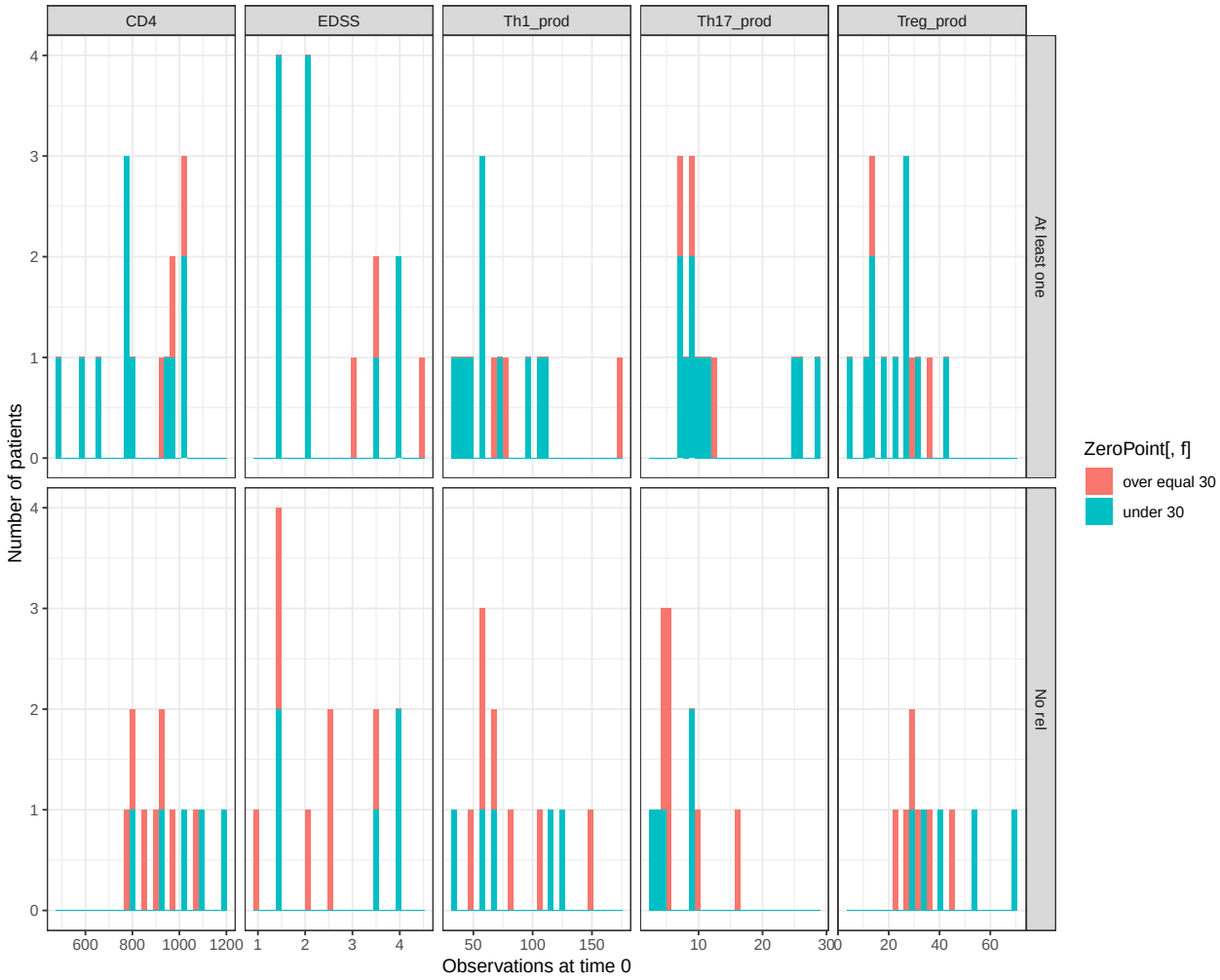
Data at the beginning time

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



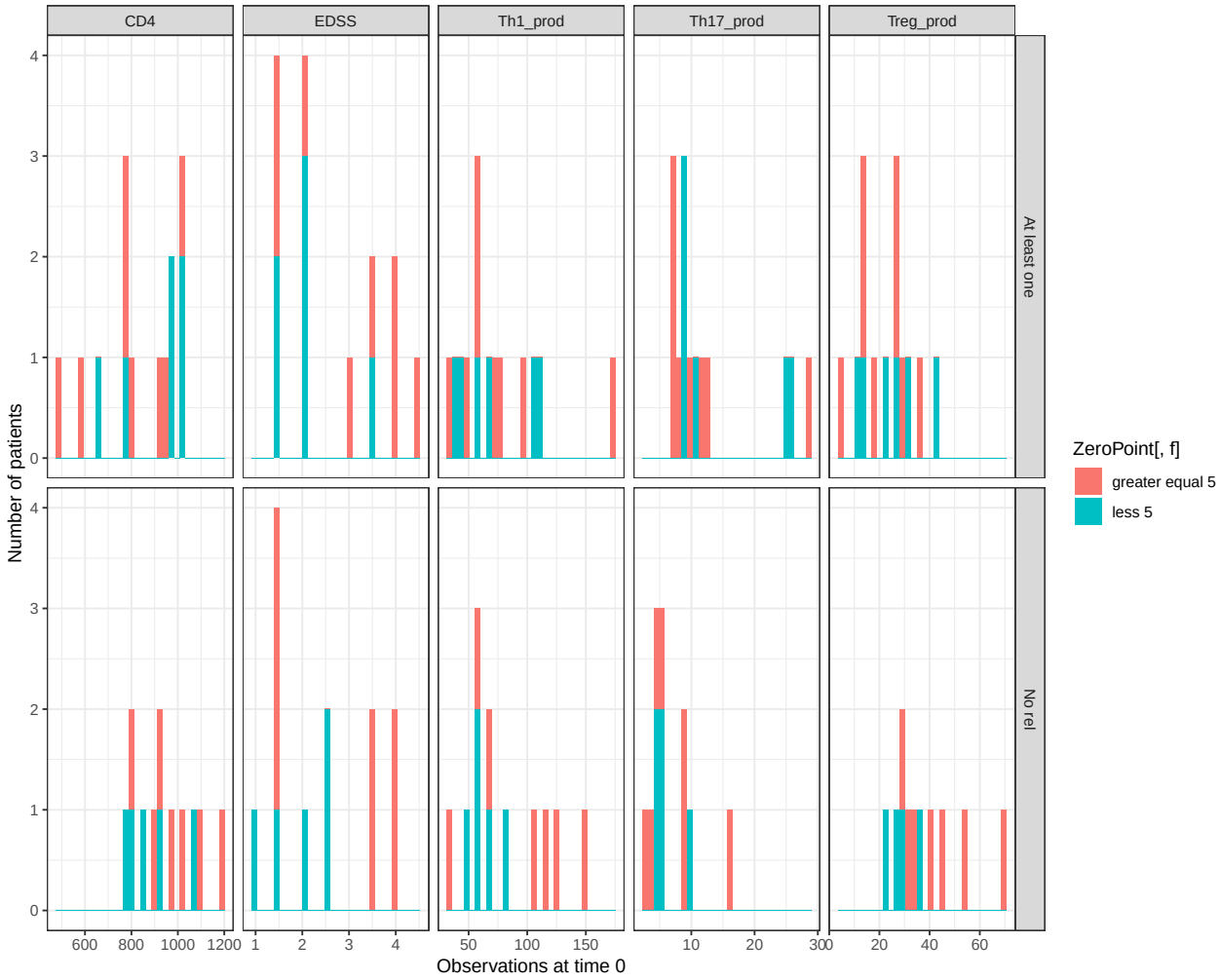
	CD4	EDSS	Th1_prod	Th17_prod	Treg_prod
F	14	14	14	14	14
M	12	12	12	12	11

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



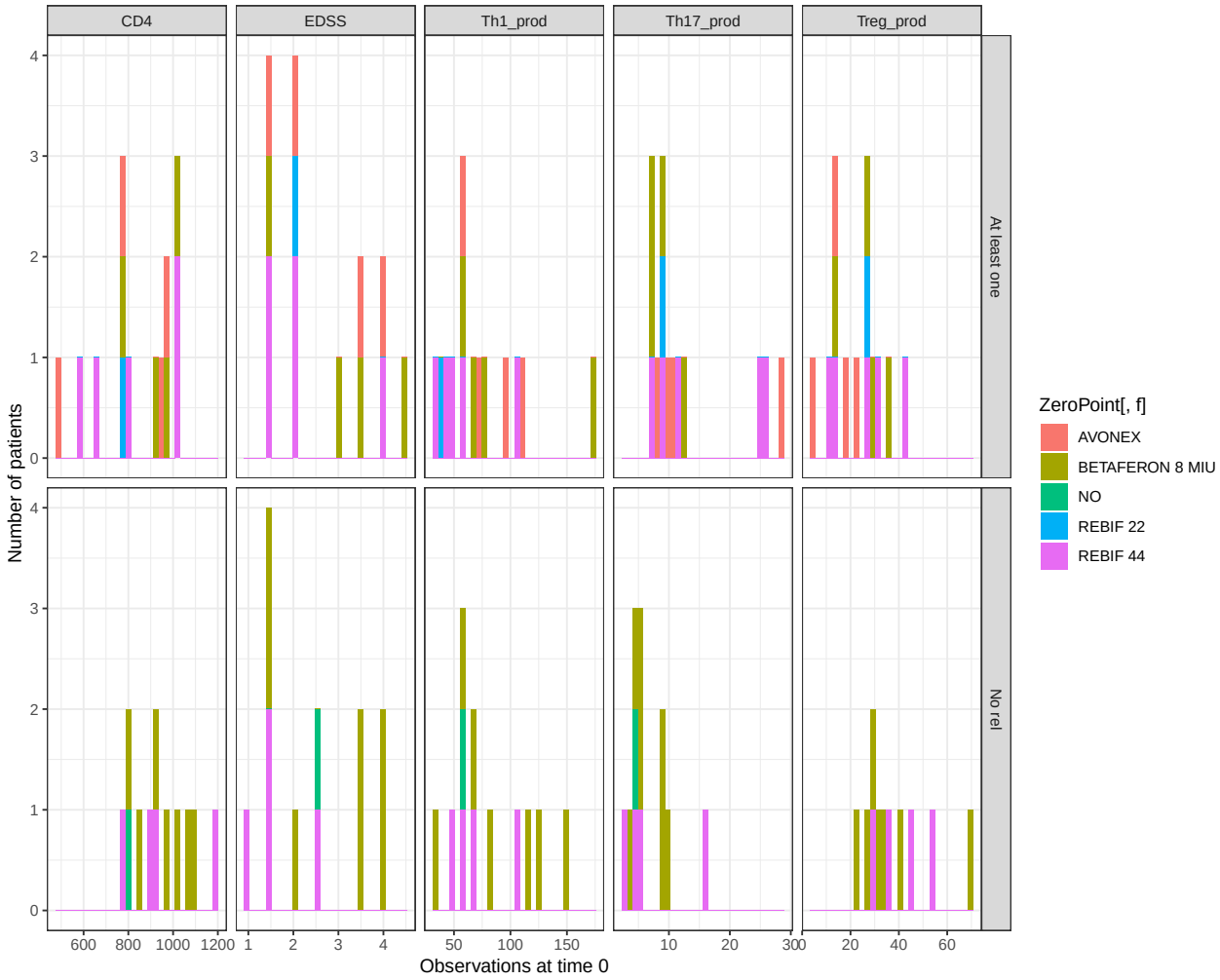
	CD4	EDSS	Th1_prod	Th17_prod	Treg_prod
over equal 30	10	10	10	10	9
under 30	16	16	16	16	16

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



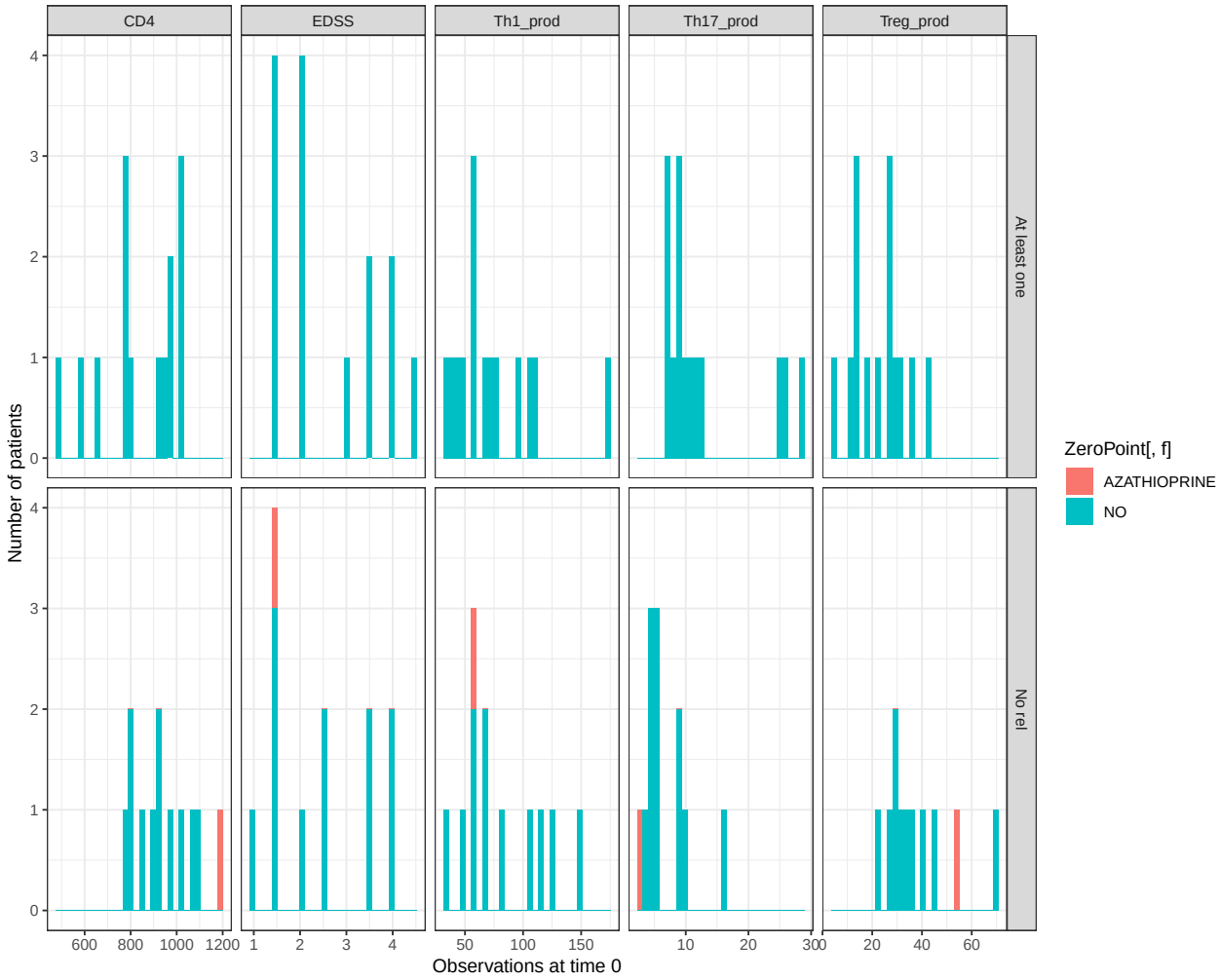
	CD4	EDSS	Th1_prod	Th17_prod	Treg_prod
greater equal 5	15	15	15	15	15
less 5	11	11	11	11	10

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



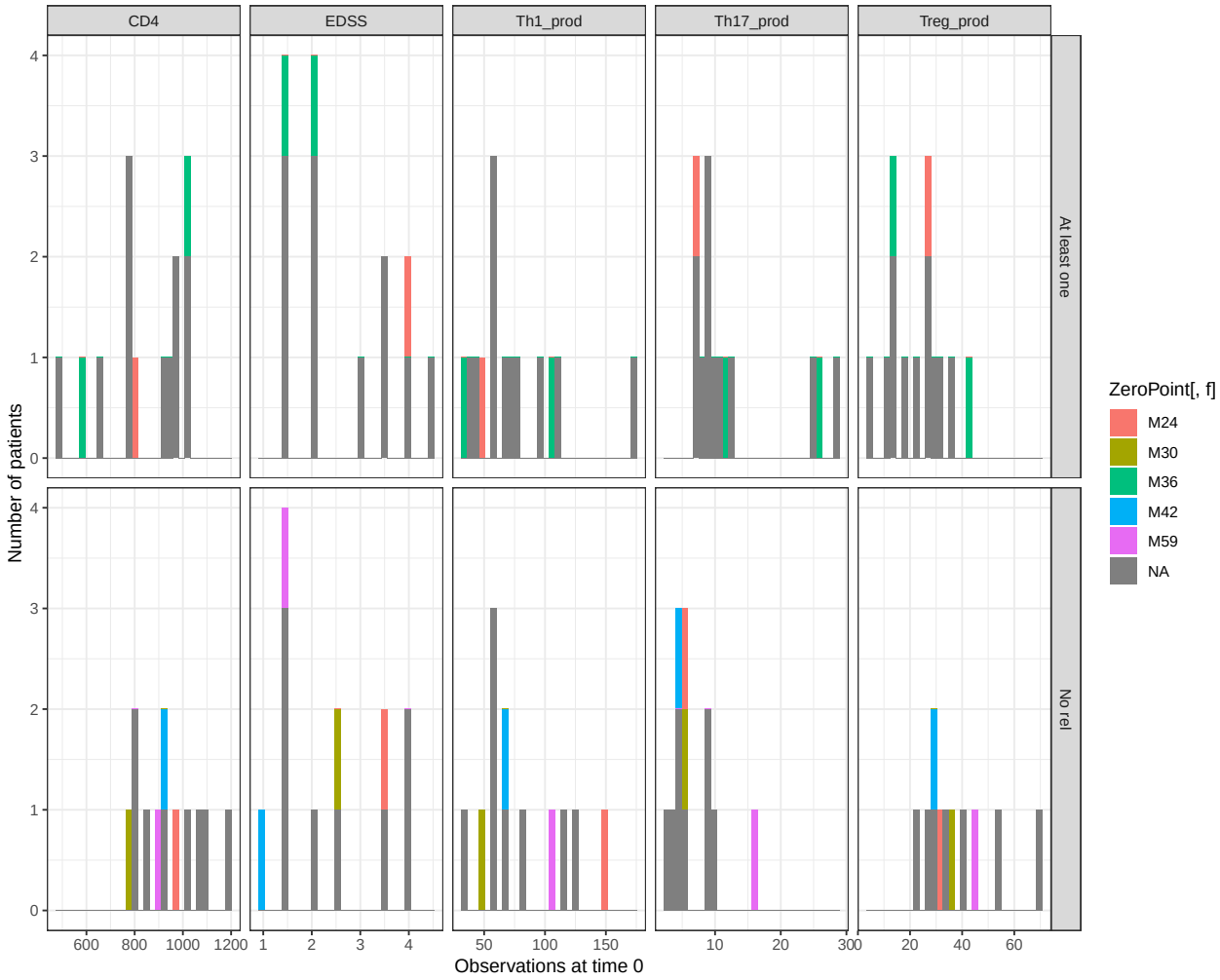
	CD4	EDSS	Th1_prod	Th17_prod	Treg_prod
AVONEX	4	4	4	4	4
BETAFERON 8 MIU	11	11	11	11	11
NO	1	1	1	1	0
REBIF 22	1	1	1	1	1
REBIF 44	9	9	9	9	9

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



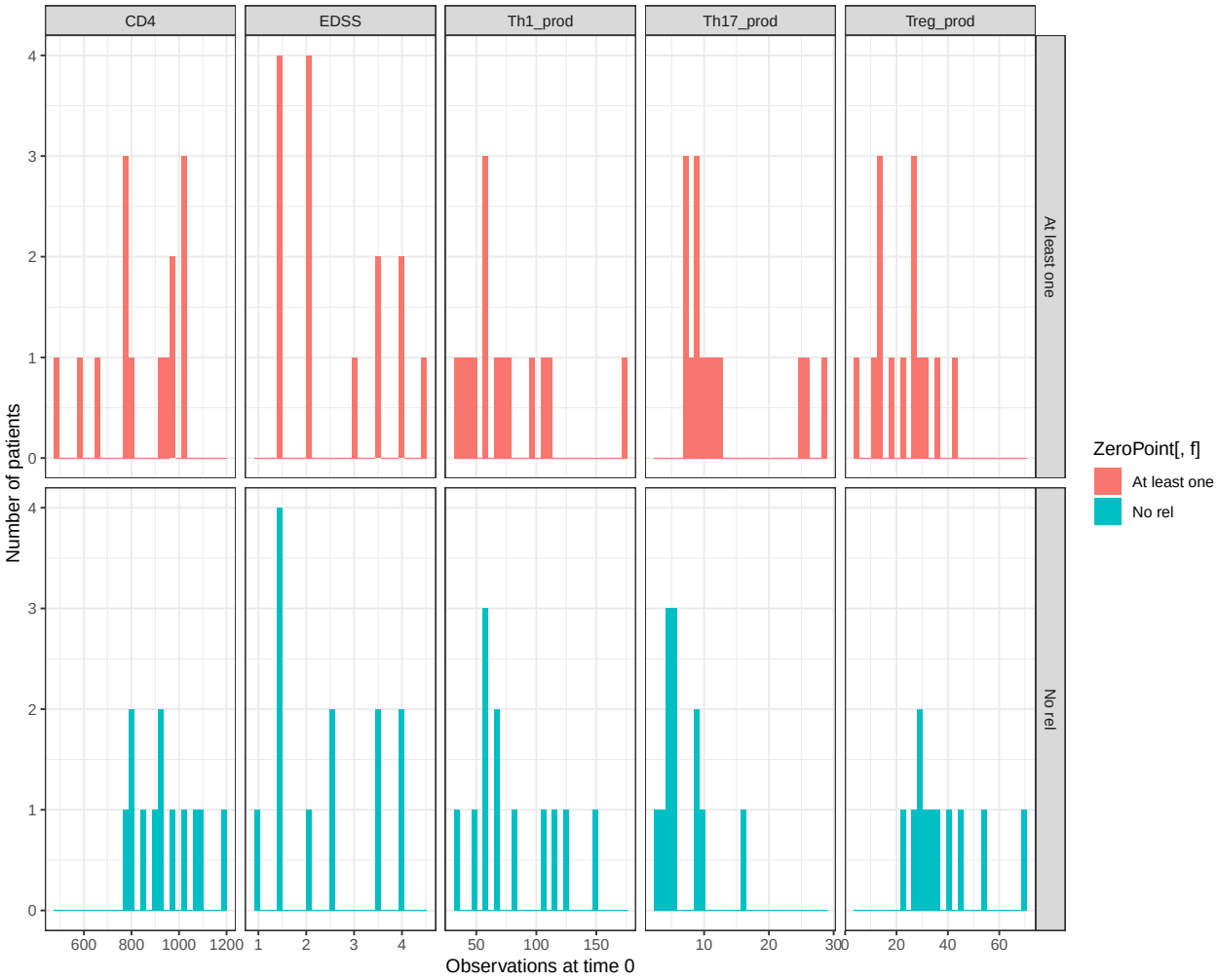
	CD4	EDSS	Th1_prod	Th17_prod	Treg_prod
AZATHIOPRINE	1	1	1	1	1
NO	25	25	25	25	24

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

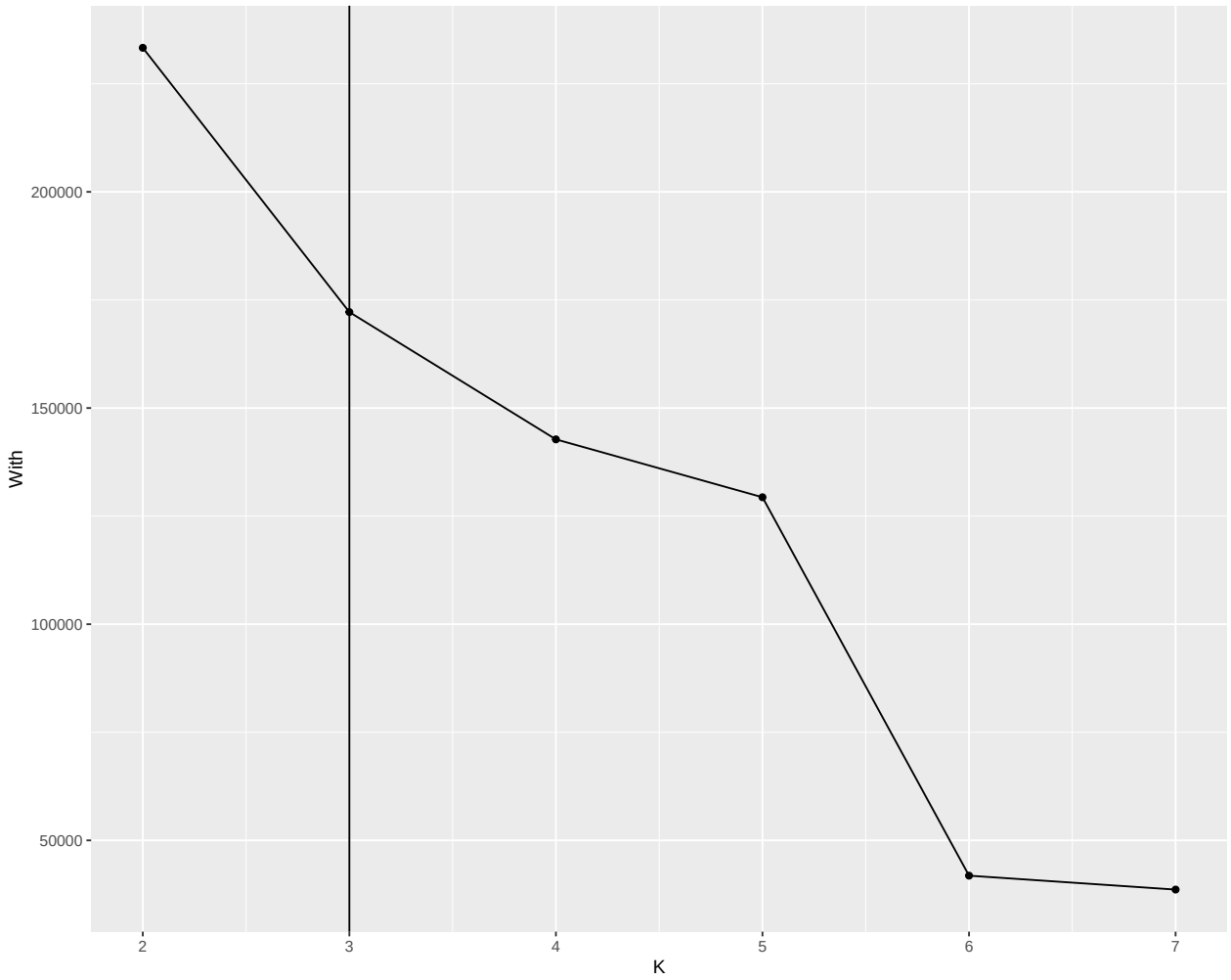


	CD4	EDSS	Th1_prod	Th17_prod	Treg_prod
M24	2	2	2	2	2
M30	1	1	1	1	1
M36	2	2	2	2	2
M42	1	1	1	1	1
M59	1	1	1	1	1

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



	CD4	EDSS	Th1_prod	Th17_prod	Treg_prod
At least one	14	14	14	14	14
No rel	12	12	12	12	11



Statistical test

GENDER

- **shapiro.test to check the normal distribution**

Type	Var1	shapiro.statistic	shapiro.p.value	Hp
Treg_prod	F	0.9748582	0.9339960	Normal
Treg_prod	M	0.8818828	0.1099269	Normal
Th17_prod	F	0.8148294	0.0076585	Not normal
Th17_prod	M	0.9365399	0.4545889	Normal
Th1_prod	F	0.9370820	0.3822134	Normal
Th1_prod	M	0.9091525	0.2080746	Normal
CD4	F	0.8485169	0.0212052	Not normal
CD4	M	0.9536749	0.6911640	Normal
EDSS	F	0.8943993	0.0935612	Normal
EDSS	M	0.8065133	0.0111214	Not normal

These have at least one group not normal distributed: Th17_prod, CD4, EDSS.

- **Wilcoxon-Mann-Whitne or t-test (if both groups are normal) to check if the groups are similar**

Type	Test	p.value	Hp
Treg_prod	M vs F	0.3114465	not different
Th17_prod	M vs F	0.0201601	different
Th1_prod	M vs F	0.0207857	different
CD4	M vs F	0.9794759	not different
EDSS	M vs F	0.4151667	not different

Discrete AGE AT ONSET

- **shapiro.test to check the normal distribution**

Type	Var1	shapiro.statistic	shapiro.p.value	Hp
Treg_prod	over equal 30	0.9758201	0.9394165	Normal
Treg_prod	under 30	0.9476682	0.4536732	Normal
Th17_prod	over equal 30	0.8485556	0.0558339	Normal
Th17_prod	under 30	0.7653409	0.0009782	Not normal
Th1_prod	over equal 30	0.8485447	0.0558174	Normal
Th1_prod	under 30	0.8946350	0.0659827	Normal
CD4	over equal 30	0.9783564	0.9557683	Normal
CD4	under 30	0.9661389	0.7728559	Normal
EDSS	over equal 30	0.9652081	0.8432146	Normal
EDSS	under 30	0.7572488	0.0007778	Not normal

These have at least one group not normal distributed: Th17_prod, EDSS.

- **Wilcoxon-Mann-Whitne or t-test (if both groups are normal) to check if the groups are similar**

Type	Test	p.value	Hp
Treg_prod	over equal 30 vs under 30	0.8743319	not different
Th17_prod	over equal 30 vs under 30	0.3913667	not different
Th1_prod	over equal 30 vs under 30	0.2680100	not different
CD4	over equal 30 vs under 30	0.3785428	not different
EDSS	over equal 30 vs under 30	0.9355991	not different

Discrete DISEASE DURATION

- **shapiro.test to check the normal distribution**

Type	Var1	shapiro.statistic	shapiro.p.value	Hp
Treg_prod	greater equal 5	0.9650367	0.7790015	Normal
Treg_prod	less 5	0.9720166	0.9088863	Normal

Type	Var1	shapiro.statistic	shapiro.p.value	Hp
Th17_prod	greater equal 5	0.7870421	0.0025190	Not normal
Th17_prod	less 5	0.7256665	0.0010031	Not normal
Th1_prod	greater equal 5	0.9325730	0.2980324	Normal
Th1_prod	less 5	0.8825464	0.1120737	Normal
CD4	greater equal 5	0.9613660	0.7161275	Normal
CD4	less 5	0.9534019	0.6876242	Normal
EDSS	greater equal 5	0.8198051	0.0066767	Not normal
EDSS	less 5	0.9150507	0.2795464	Normal

These have at least one group not normal distributed: Th17_prod, EDSS.

- **Wilcoxon-Mann-Whitne or t-test (if both groups are normal) to check if the groups are similar**

Type	Test	p.value	Hp
Treg_prod	greater equal 5 vs less 5	0.4118247	not different
Th17_prod	greater equal 5 vs less 5	0.7988051	not different
Th1_prod	greater equal 5 vs less 5	0.1788710	not different
CD4	greater equal 5 vs less 5	0.8436992	not different
EDSS	greater equal 5 vs less 5	0.0896089	not different

Relapses

- **shapiro.test to check the normal distribution**

Type	Var1	shapiro.statistic	shapiro.p.value	Hp
Treg_prod	At least one	0.9776047	0.9585388	Normal
Treg_prod	No rel	0.8679475	0.0729481	Normal
Th17_prod	At least one	0.7111602	0.0004833	Not normal
Th17_prod	No rel	0.8172960	0.0148349	Not normal
Th1_prod	At least one	0.8711723	0.0435829	Not normal
Th1_prod	No rel	0.9176973	0.2674123	Normal
CD4	At least one	0.8941724	0.0928572	Normal
CD4	No rel	0.9491245	0.6242224	Normal
EDSS	At least one	0.8493405	0.0217573	Not normal
EDSS	No rel	0.8782745	0.0833003	Normal

These have at least one group not normal distributed: Th17_prod, Th1_prod, EDSS.

- **Wilcoxon-Mann-Whitne or t-test (if both groups are normal) to check if the groups are similar**

Type	Test	p.value	Hp
Treg_prod	No rel vs At least one	0.0086082	different

Type	Test	p.value	Hp
Th17_prod	No rel vs At least one	0.0053860	different
Th1_prod	No rel vs At least one	0.7424463	not different
CD4	No rel vs At least one	0.0811755	not different
EDSS	No rel vs At least one	0.6174870	not different