

FAHMM: Roche Data

2024-12-12

Summarise the model fit

This is a markdown for replicating new MS classification using ROche data

MS dimensions

```
load('interim_tables/FA_baseline_Relapse.Rdata')
l = result_40$B
#order of latent varibales (disability,brain,relapse,Gd)
lambda=l[,c(1,2,3,4)]
rownames(lambda)=c("EDSS","T25FWM","HPT9M","PASAT","VOLT2","NBV","NUMGDT1","RELAPSE")
lambda
```

##		[,1]	[,2]	[,3]	[,4]
##	EDSS	-0.50691846	0.10691197	-0.04380997	0.34358684
##	T25FWM	-0.54753496	0.00000000	0.00000000	0.07237835
##	HPT9M	-0.48089626	-0.04392360	0.02453008	0.26242298
##	PASAT	0.11775286	0.06905098	0.00000000	-0.36239227
##	VOLT2	-0.10179645	-0.39344957	0.01926232	0.61953188
##	NBV	0.18035721	-0.04265002	0.01006890	-0.64000120
##	NUMGDT1	0.01985051	-0.61840968	0.07964227	0.07929540
##	RELAPSE	0.00000000	-0.02878093	0.38249157	0.00000000

MS dimensions: thresholded

```
load('interim_tables/FA_baseline_Relapse.Rdata')
l = result_40$B
l[result_40$P_star<0.5]=0
#order of latent varibales (disability,brain,relapse,Gd)
lambda=l[,c(1,2,3,4)]
rownames(lambda)=c("EDSS","T25FWM","HPT9M","PASAT","VOLT2","NBV","NUMGDT1","RELAPSE")

colnames(lambda)=c('disability','Asymptomatic','Relapse','brain_damage')
lambda
```

##		disability	Asymptomatic	Relapse	brain_damage
##	EDSS	-0.5069185	0.0000000	0.0000000	0.3435868
##	T25FWM	-0.5475350	0.0000000	0.0000000	0.0000000
##	HPT9M	-0.4808963	0.0000000	0.0000000	0.0000000
##	PASAT	0.0000000	0.0000000	0.0000000	-0.3623923
##	VOLT2	0.0000000	-0.3934496	0.0000000	0.6195319
##	NBV	0.0000000	0.0000000	0.0000000	-0.6400012
##	NUMGDT1	0.0000000	-0.6184097	0.0000000	0.0000000
##	RELAPSE	0.0000000	0.0000000	0.3824916	0.0000000

States Mean

```

follow      = read.csv('interim_tables/follow_FAhmm_Relapse_ALLC.csv')
follow=follow%>%filter(! USUBJID %in% 'WA21093-208650-1932631')
follow$VOLT2 = follow$VOLT2^3
yy          = as.matrix(follow[,c("V32","V33","V34","V35")])
yy          = scale(yy,center = FALSE,scale = TRUE)

follow <- follow %>%
  mutate(MONTHN = MONTH,
         MONTHN = ifelse(MONTHN == -1, 0, MONTHN)) %>%
  group_by(USUBJID) %>%
  arrange(MONTHN, .by_group = TRUE) %>%
  mutate(deltaM = -lag(MONTHN) + MONTHN) %>%
  ungroup() %>%
  mutate(deltaM = ifelse(is.na(deltaM), 0, deltaM))

Time      = follow$deltaM
ss        = follow%>%group_by(USUBJID)%>%summarise(s=length(USUBJID))
seq=ss$s

load('interim_tables/hmm_s8_init_clara2_euc.RData')
hmm_CTDC = hmm_CTDC_clara2_euc
l = order(hmm_CTDC$mu[1,],decreasing = TRUE)
hmm_CTDC$mu = hmm_CTDC$mu[,l]
hmm_CTDC$A  = hmm_CTDC$A[l,l]
hmm_CTDC$pi = hmm_CTDC$pi[l]
hmm_CTDC$sigma = hmm_CTDC$sigma[, ,l]

Z = ctdthmm_MultSubj_viterbi(hmm_CTDC,yy,seq,Time)
follow$Z=Z
table(Z)

## Z
##      1      2      3      4      5      6      7      8
## 7166 6871 5592 1190  963 3410 5666 3241

follow$Z=Z
follow$MS=NA
follow$MS[Z%in%c(1,2,3)]=1
follow$MS[Z==4]=2
follow$MS[Z==5]=3
follow$MS[Z%in%c(6,7,8)]=4

write.csv(follow, 'interim_tables/follow_FAhmm_Relapse_ALLC_MetaS.csv')
kk= follow%>%dplyr::select(c( "EDSS", "T25FWM", "HPT9M", "PASAT", "VOLT2", "NBV", "NUMGDT1", "RELAPSE"))
spl=split(kk,Z)
Emp_mu_S9 = sapply(spl,function(x){colMeans(x,na.rm = TRUE)})
Emp_mu_S9[6,]=Emp_mu_S9[6,]/1000

```

States mean original

```
round(Emp_mu_S9,2)
```

```
##           1      2      3      4      5      6      7      8
## EDSS      2.59  3.03  3.32  3.02  3.85  4.65  6.02  6.55
## T25FWM    5.72  6.62  7.07 10.19 12.57 12.31 17.98 113.40
## HPT9M     20.73 23.17 25.22 26.14 28.97 31.11 31.43 89.49
## PASAT     52.67 48.77 44.15 42.85 45.73 36.25 48.10 42.45
## VOLT2      1.37  5.46 14.90 14.11 12.54 37.17  7.06 16.71
## NBV        1.54  1.49  1.45  1.50  1.50  1.40  1.46  1.41
## NUMGDT1    0.00  0.00  0.00  3.64  1.59  0.11  0.00  0.43
## RELAPSE    0.00  0.00  0.00  0.00  1.00  0.00  0.00  0.00
```

States mean composite

```
mu= hmm_CTDC$mu
row.names(mu)=c('disability','Asymptomatic','Relapse','brain_damage')
mu
```

```
##           [,1]      [,2]      [,3]      [,4]      [,5]
## disability    0.4031489  0.3463907  0.3259606  0.2790610  0.03398707
## Asymptomatic  0.5211825  0.1560141 -0.3373005 -2.7013025 -0.37500438
## Relapse       -0.1790812 -0.1872687 -0.1794779  0.2438864  5.74551639
## brain_damage -1.0119571 -0.1561804  0.5657804 -0.1398755 -0.06598466
##           [,6]      [,7]      [,8]
## disability    0.02172144 -0.4539751 -2.45173448
## Asymptomatic -0.86098357  0.6123441  0.04840896
## Relapse       -0.21165852 -0.3545487 -0.15403517
## brain_damage  1.55302596  0.1924550  0.61013640
```

Transition probability

```
AD=hmm_CTDC$A
AD[AD<.005]=0
round(AD,4)
```

```
##           [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]
## [1,] 0.9560 0.0000 0.0000 0.0112 0.0234 0.0000 0.0060 0.0000
## [2,] 0.0000 0.9446 0.0000 0.0177 0.0210 0.0000 0.0105 0.0000
## [3,] 0.0000 0.0000 0.9343 0.0281 0.0203 0.0000 0.0096 0.0000
## [4,] 0.1639 0.2636 0.3065 0.0420 0.0150 0.1297 0.0542 0.0252
## [5,] 0.1870 0.1412 0.1129 0.0326 0.3631 0.0906 0.0484 0.0243
## [6,] 0.0000 0.0000 0.0000 0.0145 0.0246 0.9410 0.0000 0.0164
## [7,] 0.0000 0.0000 0.0000 0.0000 0.0060 0.0000 0.9433 0.0421
## [8,] 0.0000 0.0000 0.0000 0.0000 0.0066 0.0054 0.0432 0.9371
```

Baseline classification

```
pi=hmm_CTDC$pi
round(pi,2)
```

```
## [1] 0.21 0.16 0.10 0.34 0.04 0.06 0.08 0.02
```

Demographic Summary

```
follow$Z=Z
follow$MS=NA
follow$MS[Z%in%c(1,2,3)]=1
follow$MS[Z%in%c(6,7,8)]=4
follow$MS[Z==4]=4
follow$MS[Z==5]=5
ff2<-follow%>%dplyr::select(c("AGELG","SEX","MSTYPE","DURFSLG","RELST1Y","Z"))%>%tbl_summary(by=Z, mis=
label = 1,
type = li
statistic

ff2
```

```
## Table printed with `knitr::kable()`, not {gt}. Learn why at
## https://www.danielsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include `message = FALSE` in code chunk header.
```

Variable	1, N = 7,166	2, N = 6,871	3, N = 5,592	4, N = 1,190	5, N = 963	6, N = 3,410	7, N = 5,666	8, N = 3,241
Age								
Mean (SD)	41 (10)	43 (10)	42 (9)	36 (9)	38 (9)	45 (10)	48 (8)	47 (9)
Median (IQR)	42 (33, 50)	43 (35, 51)	44 (36, 50)	36 (29, 43)	38 (31, 45)	46 (38, 52)	49 (43, 54)	49 (41, 54)
Gender								
F	4,565 (64%)	3,732 (54%)	3,085 (55%)	694 (58%)	645 (67%)	2,040 (60%)	2,936 (52%)	1,490 (46%)
M	2,601 (36%)	3,139 (46%)	2,507 (45%)	496 (42%)	318 (33%)	1,370 (40%)	2,730 (48%)	1,751 (54%)
MS Type								
PPMS	2,629 (37%)	3,175 (46%)	2,852 (51%)	291 (24%)	135 (14%)	2,194 (64%)	5,227 (92%)	3,096 (96%)
RMS	4,537 (63%)	3,696 (54%)	2,740 (49%)	899 (76%)	828 (86%)	1,216 (36%)	439 (7.7%)	145 (4.5%)
Duration from First Symptom								
Mean (SD)	6.9 (4.8)	7.9 (5.1)	9.3 (5.5)	6.2 (5.3)	7.3 (5.9)	10.4 (5.6)	10.8 (5.0)	12.1 (4.7)
Median (IQR)	6.1 (3.0, 9.5)	7.2 (3.8, 10.9)	8.7 (5.2, 12.3)	4.9 (2.3, 8.7)	5.8 (2.7, 10.5)	9.6 (6.3, 13.5)	10.1 (7.3, 13.4)	11.7 (8.7, 15.0)
Number of Relapses during past year								
Mean (SD)	1 (1)	1 (1)	1 (1)	1 (1)	1 (1)	1 (1)	0 (0)	0 (0)
Median (IQR)	1 (0, 1)	1 (0, 1)	0 (0, 1)	1 (0, 1)	1 (1, 2)	0 (0, 1)	0 (0, 0)	0 (0, 0)

Features Summary

```
ff3<-follow%>%dplyr::select(c("MSTYPE","EDSS","T25FWM","HPT9M","PASAT","VOLT2","NBV","NUMGDT1","RELAPSE
```

ff3

Table printed with `knitr::kable()`, not {gt}. Learn why at
 ## <https://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html>
 ## To suppress this message, include `message = FALSE` in code chunk header.

Variable	1, N = 7,166	2, N = 6,871	3, N = 5,592	4, N = 1,190	5, N = 963	6, N = 3,410	7, N = 5,666	8, N = 3,241
MSTYPE								
PPMS	2,629 (37%)	3,175 (46%)	2,852 (51%)	291 (24%)	135 (14%)	2,194 (64%)	5,227 (92%)	3,096 (96%)
RMS	4,537 (63%)	3,696 (54%)	2,740 (49%)	899 (76%)	828 (86%)	1,216 (36%)	439 (7.7%)	145 (4.5%)
EDSS								
Mean	2.59	3.03	3.32	3.02	3.85	4.65	6.02	6.55
(SD)	(1.27)	(1.23)	(1.08)	(1.48)	(1.54)	(1.46)	(0.43)	(1.21)
Median	2.50	3.00	3.50	3.00	4.00	4.50	6.00	6.50
(IQR)	(1.50, 3.50)	(2.00, 4.00)	(2.50, 4.00)	(2.00, 4.00)	(2.50, 5.00)	(3.50, 6.00)	(6.00, 6.50)	(6.50, 7.00)
(10%, 90%)	(1.00, 4.00)	(1.50, 4.50)	(2.00, 4.50)	(1.50, 5.50)	(2.00, 6.00)	(2.50, 6.50)	(5.50, 6.50)	(4.50, 8.00)
Walking Time								
Mean	6 (2)	7 (3)	7 (4)	10 (13)	13 (27)	12 (11)	18 (11)	113 (70)
(SD)								
Median	5 (4, 6)	6 (5, 7)	6 (5, 8)	6 (5, 9)	6 (5, 9)	9 (6, 14)	15 (10, 22)	136 (44, 180)
(IQR)								
(10%, 90%)	(4, 8)	(4, 10)	(4, 10)	(4, 19)	(4, 16)	(5, 24)	(8, 33)	(15, 180)
Hand Coordina- tion								
Mean	21 (4)	23 (5)	25 (6)	26 (10)	29 (30)	31 (9)	31 (11)	89 (80)
(SD)								
Median	20 (18, 23)	22 (20, 26)	24 (21, 28)	23 (20, 29)	23 (20, 29)	30 (24, 36)	29 (23, 37)	59 (39, 100)
(IQR)								
(10%, 90%)	(16, 26)	(18, 30)	(19, 33)	(18, 38)	(18, 36)	(21, 43)	(20, 48)	(27, 253)
PASAT								
Mean	53 (9)	49 (11)	44 (13)	43 (13)	46 (13)	36 (14)	48 (12)	42 (14)
(SD)								
Median	56 (50, 59)	53 (43, 58)	48 (34, 55)	46 (34, 54)	50 (38, 57)	36 (26, 48)	52 (41, 58)	45 (30, 56)
(IQR)								
(10%, 90%)	(40, 60)	(31, 60)	(25, 59)	(25, 58)	(26, 59)	(18, 56)	(29, 60)	(22, 59)
T2 lesion volume								
Mean	1 (1)	5 (2)	15 (5)	14 (15)	13 (15)	37 (14)	7 (7)	17 (18)
(SD)								
Median	1 (1, 2)	5 (4, 7)	14 (11, 18)	10 (4, 18)	7 (2, 17)	34 (27, 45)	5 (2, 10)	10 (3, 26)
(IQR)								

Variable	1, N = 7,166	2, N = 6,871	3, N = 5,592	4, N = 1,190	5, N = 963	6, N = 3,410	7, N = 5,666	8, N = 3,241
(10%, 90%)	(0, 3)	(3, 8)	(10, 21)	(2, 33)	(1, 32)	(22, 56)	(1, 17)	(1, 45)
Normalised brain volume								
Mean	1,542	1,490	1,448	1,498	1,499	1,395	1,461	1,411
(SD)	(70)	(76)	(80)	(90)	(93)	(82)	(73)	(74)
Median	1,543	1,491	1,446	1,497	1,503	1,389	1,460	1,406
(IQR)	(1,492, 1,589)	(1,437, 1,543)	(1,389, 1,505)	(1,437, 1,561)	(1,439, 1,568)	(1,341, 1,459)	(1,415, 1,506)	(1,372, 1,449)
(10%, 90%)	(1,450, 1,631)	(1,390, 1,592)	(1,340, 1,554)	(1,385, 1,611)	(1,371, 1,610)	(1,290, 1,497)	(1,370, 1,553)	(1,325, 1,513)
Number of T1 Gd lesions								
Mean	0 (0)	0 (0)	0 (0)	4 (6)	2 (5)	0 (0)	0 (0)	0 (4)
(SD)								
Median	0 (0, 0)	0 (0, 0)	0 (0, 0)	2 (1, 4)	0 (0, 1)	0 (0, 0)	0 (0, 0)	0 (0, 0)
(IQR)								
(10%, 90%)	(0, 0)	(0, 0)	(0, 0)	(1, 9)	(0, 4)	(0, 0)	(0, 0)	(0, 0)
RELAPSE	0 (0%)	0 (0%)	0 (0%)	0 (0%)	963 (100%)	0 (0%)	0 (0%)	0 (0%)

Demographic Summary Meta state

```
follow$Z=Z
follow$MS=NA
follow$MS[Z%in%c(1,2,3)]=1
follow$MS[Z==4]=2
follow$MS[Z==5]=3
follow$MS[Z%in%c(6,7,8)]=4
```

```
ff2<-follow%>%dplyr::select(c("AGELG", "SEX", "MSTYPE", "DURFSLG", "REL PST1Y", "MS"))%>%tbl_summary(by=MS, m
label = 1,
type = li
statistic
```

```
## Table printed with `knitr::kable()`, not {gt}. Learn why at
## https://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include `message = FALSE` in code chunk header.
```

Variable	1, N = 19,629	2, N = 1,190	3, N = 963	4, N = 12,317
Age				
Mean (SD)	42 (10)	36 (9)	38 (9)	47 (9)
Median (IQR)	43 (35, 50)	36 (29, 43)	38 (31, 45)	48 (41, 54)
Gender				
F	11,382 (58%)	694 (58%)	645 (67%)	6,466 (52%)
M	8,247 (42%)	496 (42%)	318 (33%)	5,851 (48%)

Variable	1, N = 19,629	2, N = 1,190	3, N = 963	4, N = 12,317
MS Type				
PPMS	8,656 (44%)	291 (24%)	135 (14%)	10,517 (85%)
RMS	10,973 (56%)	899 (76%)	828 (86%)	1,800 (15%)
Duration from First Symptom				
Mean (SD)	7.9 (5.2)	6.2 (5.3)	7.3 (5.9)	11.0 (5.1)
Median (IQR)	7.2 (3.7, 10.9)	4.9 (2.3, 8.7)	5.8 (2.7, 10.5)	10.4 (7.4, 14.0)
Number of Relapses during past year				
Mean (SD)	1 (1)	1 (1)	1 (1)	0 (1)
Median (IQR)	1 (0, 1)	1 (0, 1)	1 (1, 2)	0 (0, 0)

Features Summary Meta State

```
ff3<-follow%>%dplyr::select(c("MSTYPE","EDSS","T25FWM","HPT9M","PASAT","VOLT2","NBV","NUMGDT1","RELAPSE
```

```
ff3
```

```
## Table printed with `knitr::kable()`, not {gt}. Learn why at
## https://www.danielsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include `message = FALSE` in code chunk header.
```

Variable	1, N = 19,629	2, N = 1,190	3, N = 963	4, N = 12,317
MSTYPE				
PPMS	8,656 (44%)	291 (24%)	135 (14%)	10,517 (85%)
RMS	10,973 (56%)	899 (76%)	828 (86%)	1,800 (15%)
EDSS				
Mean (SD)	2.95 (1.24)	3.02 (1.48)	3.85 (1.54)	5.78 (1.26)
Median (IQR)	3.00 (2.00, 4.00)	3.00 (2.00, 4.00)	4.00 (2.50, 5.00)	6.00 (5.50, 6.50)
(10%, 90%)	(1.50, 4.50)	(1.50, 5.50)	(2.00, 6.00)	(4.00, 7.00)
Walking Time				
Mean (SD)	6 (3)	10 (13)	13 (27)	41 (56)
Median (IQR)	6 (5, 7)	6 (5, 9)	6 (5, 9)	16 (9, 36)
(10%, 90%)	(4, 9)	(4, 19)	(4, 16)	(7, 180)
Hand Coordination				
Mean (SD)	23 (5)	26 (10)	29 (30)	46 (49)
Median (IQR)	22 (19, 26)	23 (20, 29)	23 (20, 29)	32 (25, 45)
(10%, 90%)	(17, 30)	(18, 38)	(18, 36)	(21, 73)
PASAT				
Mean (SD)	49 (12)	43 (13)	46 (13)	43 (14)
Median (IQR)	53 (43, 58)	46 (34, 54)	50 (38, 57)	47 (32, 56)
(10%, 90%)	(30, 60)	(25, 58)	(26, 59)	(23, 59)
T2 lesion volume				
Mean (SD)	6 (6)	14 (15)	13 (15)	18 (18)
Median (IQR)	4 (2, 10)	10 (4, 18)	7 (2, 17)	12 (3, 30)
(10%, 90%)	(1, 16)	(2, 33)	(1, 32)	(1, 45)
Normalised brain volume				
Mean (SD)	1,502 (84)	1,498 (90)	1,499 (93)	1,421 (84)

Variable	1, N = 19,629	2, N = 1,190	3, N = 963	4, N = 12,317
Median (IQR)	1,505 (1,443, 1,562)	1,497 (1,437, 1,561)	1,503 (1,439, 1,568)	1,424 (1,362, 1,476)
(10%, 90%)	(1,390, 1,610)	(1,385, 1,611)	(1,371, 1,610)	(1,317, 1,519)
Number of T1 Gd lesions				
Mean (SD)	0 (0)	4 (6)	2 (5)	0 (2)
Median (IQR)	0 (0, 0)	2 (1, 4)	0 (0, 1)	0 (0, 0)
(10%, 90%)	(0, 0)	(1, 9)	(0, 4)	(0, 0)
RELAPSE	0 (0%)	0 (0%)	963 (100%)	0 (0%)

States density plots: baseline original features

```
p1<-ggplot(follow, aes(y=as.factor(Z), x=EDSS, fill=as.factor(Z))) +
  geom_density_ridges(alpha=0.6, bandwidth=1) +
  scale_fill_viridis(discrete=TRUE) +
  scale_color_viridis(discrete=TRUE) +
  theme_bw() +
  theme(
    legend.position="none",
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  xlab("EDSS (total score)") +
  ylab("States")

p2<-ggplot(follow, aes(y=as.factor(Z), x=T25FWM, fill=as.factor(Z))) +
  geom_density_ridges(alpha=0.6, bandwidth=4) +xlim(0,50)+
  scale_fill_viridis(discrete=TRUE) +
  scale_color_viridis(discrete=TRUE) +
  theme_bw() +
  theme(
    legend.position="none",
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  xlab("25-foot times walking test (s)") +
  ylab("")

p3<-ggplot(follow, aes(y=as.factor(Z), x=HPT9M, fill=as.factor(Z))) +
  geom_density_ridges(alpha=0.6, bandwidth=4) +xlim(0,100)+
  scale_fill_viridis(discrete=TRUE) +
  scale_color_viridis(discrete=TRUE) +
  theme_bw() +
  theme(
    legend.position="none",
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  xlab("Hand Coordination (s)") +
  ylab("")
```



```

p4<-ggplot(follow, aes(y=as.factor(Z), x=PASAT, fill=as.factor(Z))) +
  geom_density_ridges(alpha=0.6, bandwidth=3) +xlim(0,60)+
  scale_fill_viridis(discrete=TRUE) +
  scale_color_viridis(discrete=TRUE) +
  theme_bw() +
  theme(
    legend.position="none",
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  xlab("PASAT (correct answers out of 60)") +
  ylab("")

p5<-ggplot(follow, aes(y=as.factor(Z), x=NBV/1000, fill=as.factor(Z))) +
  geom_density_ridges(alpha=0.6, bandwidth=.05) +
  scale_fill_viridis(discrete=TRUE) +
  scale_color_viridis(discrete=TRUE) +
  theme_bw() +
  theme(
    legend.position="none",
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  xlab("Normalised Brain Volume (L)") +
  ylab("States")

p6<-ggplot(follow, aes(y=as.factor(Z), x=VOLT2, fill=as.factor(Z))) +
  geom_density_ridges(alpha=0.6, bandwidth=.3) +xlim(NA,5)+
  scale_fill_viridis(discrete=TRUE) +
  scale_color_viridis(discrete=TRUE) +
  theme_bw() +
  theme(
    legend.position="none",
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  xlab("T2 Lesion Volume") +
  ylab("")

p7<-ggplot(follow, aes(y=as.factor(Z), x=NUMGDT1, group=as.factor(Z))) +
  geom_density_ridges(aes(fill = as.factor(Z)), stat = "binline", binwidth = .5, scale = 0.95) +xlim(-1
  scale_fill_viridis(discrete=TRUE) +
  scale_color_viridis(discrete=TRUE) +
  theme_bw() +
  theme(
    legend.position="none",
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  xlab("Number of T1 Gd lesions") +
  ylab("")

p8<-ggplot(follow, aes(y=as.factor(Z), x=RELAPSE, group=as.factor(Z))) +

```

```

geom_density_ridges(aes(fill = as.factor(Z)), stat = "binline", binwidth = .5, scale = 0.95) +xlim(-1
scale_fill_viridis(discrete=TRUE) +
scale_color_viridis(discrete=TRUE) +
theme_bw() +
theme(
  legend.position="none",
  panel.spacing = unit(0.1, "lines"),
  strip.text.x = element_text(size = 8)
) +
xlab("Relapse Occurence") +
ylab("")

```

```

(p1|p2|p3|p4)/(p5|p6|p7|p8)

```

```

## Warning: Removed 5163 rows containing non-finite outside the scale range
## (`stat_density_ridges()`).

```

```

## Warning: Removed 3492 rows containing non-finite outside the scale range
## (`stat_density_ridges()`).

```

```

## Warning: Removed 3346 rows containing non-finite outside the scale range
## (`stat_density_ridges()`).

```

```

## Warning: Removed 28111 rows containing non-finite outside the scale range
## (`stat_density_ridges()`).

```

```

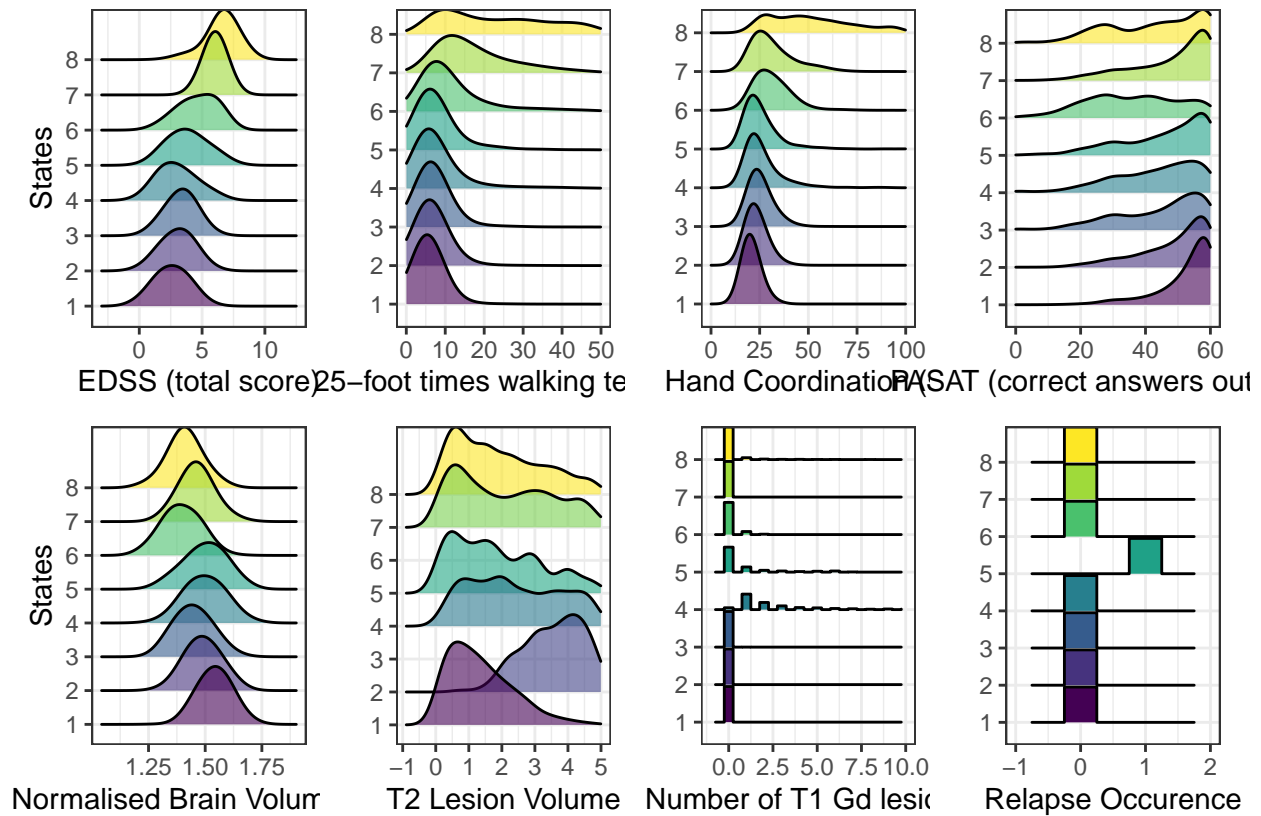
## Warning: Removed 29082 rows containing non-finite outside the scale range
## (`stat_density_ridges()`).

```

```

## Warning: Removed 23176 rows containing non-finite outside the scale range
## (`stat_binline()`).

```



Demographic Summary: Baseline

```
baseline = follow%>%filter(DAY==1)
ff2<-baseline%>%dplyr::select(c("AGE", "SEX", "MSTYPE", "DURFS", "RELPST1Y", "Z"))%>%tbl_summary(by=Z, missing=
label = 1,
type = list,
statistic = list)

ff2
```

```
## Table printed with `knitr::kable()`, not {gt}. Learn why at
## https://www.danielsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include `message = FALSE` in code chunk header.
```

Variable	1, N = 478	2, N = 354	3, N = 227	4, N = 739	5, N = 91	6, N = 133	7, N = 179	8, N = 41
Age								
Mean (SD)	39 (9)	42 (9)	42 (8)	36 (9)	37 (9)	44 (8)	46 (7)	41 (9)
Median (IQR)	39 (32, 45)	42 (36, 50)	43 (36, 49)	36 (29, 43)	38 (31, 45)	46 (39, 51)	47 (41, 51)	42 (36, 49)
Gender								
F	319 (67%)	204 (58%)	134 (59%)	435 (59%)	66 (73%)	76 (57%)	102 (57%)	21 (51%)
M	159 (33%)	150 (42%)	93 (41%)	304 (41%)	25 (27%)	57 (43%)	77 (43%)	20 (49%)
MS Type								
PPMS	100 (21%)	108 (31%)	82 (36%)	176 (24%)	8 (8.8%)	65 (49%)	141 (79%)	30 (73%)

Variable	1, N = 478	2, N = 354	3, N = 227	4, N = 739	5, N = 91	6, N = 133	7, N = 179	8, N = 41
RMS	378 (79%)	246 (69%)	145 (64%)	563 (76%)	83 (91%)	68 (51%)	38 (21%)	11 (27%)
Duration from First Symptom								
Mean (SD)	5.1 (4.8)	6.8 (5.7)	8.4 (6.4)	5.9 (5.2)	6.1 (6.5)	8.9 (5.8)	8.4 (5.0)	9.2 (4.5)
Median (IQR)	3.8 (1.6, 7.2)	5.8 (2.6, 9.5)	6.9 (3.7, 11.0)	4.5 (1.8, 8.5)	3.9 (1.5, 8.3)	7.4 (4.4, 12.8)	7.0 (5.2, 10.9)	8.6 (5.7, 12.4)
Number of Relapses during past year								
Mean (SD)	1 (1)	1 (1)	1 (1)	1 (1)	1 (1)	1 (1)	0 (1)	0 (1)
Median (IQR)	1 (1, 1)	1 (0, 1)	1 (0, 1)	1 (0, 2)	1 (1, 2)	0 (0, 1)	0 (0, 0)	0 (0, 1)

Features Summary: Baseline

```
ff3<-baseline%>%dplyr::select(c("MSTYPE", "EDSS", "T25FWM", "HPT9M", "PASAT", "VOLT2", "NBV", "NUMGDT1", "RELAPSE"))
```

```
ff3
```

```
## Table printed with `knitr::kable()`, not {gt}. Learn why at
## https://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include `message = FALSE` in code chunk header.
```

Variable	1, N = 478	2, N = 354	3, N = 227	4, N = 739	5, N = 91	6, N = 133	7, N = 179	8, N = 41
MSTYPE								
PPMS	100 (21%)	108 (31%)	82 (36%)	176 (24%)	8 (8.8%)	65 (49%)	141 (79%)	30 (73%)
RMS	378 (79%)	246 (69%)	145 (64%)	563 (76%)	83 (91%)	68 (51%)	38 (21%)	11 (27%)
EDSS								
Mean	2.56	3.07	3.36	3.13	2.86	4.45	5.78	5.87
(SD)	(1.23)	(1.14)	(1.06)	(1.51)	(1.34)	(1.30)	(0.50)	(0.97)
Median	2.50	3.00	3.50	3.00	2.50	4.50	6.00	6.50
(IQR)	(1.50, 3.50)	(2.00, 4.00)	(2.75, 4.00)	(2.00, 4.00)	(2.00, 4.00)	(3.50, 5.50)	(5.50, 6.00)	(5.50, 6.50)
(10%, 90%)	(1.00, 4.00)	(1.50, 4.50)	(2.00, 4.50)	(1.50, 5.50)	(1.50, 4.50)	(3.00, 6.00)	(5.00, 6.50)	(4.00, 6.50)
Walking Time								
Mean	6 (2)	6 (3)	7 (4)	9 (10)	7 (8)	11 (9)	17 (10)	73 (52)
(SD)								
Median	5 (4, 6)	6 (5, 7)	6 (5, 8)	6 (5, 9)	5 (4, 7)	8 (6, 12)	14 (10, 23)	58 (30, 115)
(IQR)								
(10%, 90%)	(4, 8)	(4, 10)	(4, 11)	(4, 16)	(4, 9)	(5, 19)	(8, 32)	(16, 145)

Variable	1, N = 478	2, N = 354	3, N = 227	4, N = 739	5, N = 91	6, N = 133	7, N = 179	8, N = 41
Hand Coordination								
Mean (SD)	21 (4)	24 (5)	26 (6)	26 (9)	28 (30)	31 (9)	32 (10)	82 (63)
Median (IQR)	21 (18, 23)	23 (20, 26)	25 (22, 29)	24 (20, 29)	23 (20, 28)	29 (25, 35)	29 (24, 37)	66 (45, 89)
(10%, 90%)	(17, 26)	(18, 30)	(19, 33)	(18, 37)	(18, 32)	(21, 40)	(22, 45)	(32, 138)
PASAT								
Mean (SD)	47 (11)	43 (12)	37 (14)	41 (13)	42 (13)	32 (14)	41 (13)	33 (15)
Median (IQR)	50 (41, 55)	45 (35, 53)	38 (26, 49)	43 (32, 51)	47 (30, 52)	29 (22, 41)	42 (31, 52)	30 (22, 44)
(10%, 90%)	(31, 58)	(26, 57)	(19, 55)	(23, 57)	(24, 56)	(16, 54)	(24, 56)	(16, 57)
T2 lesion volume								
Mean (SD)	1 (1)	5 (2)	15 (5)	14 (15)	12 (14)	37 (14)	6 (6)	21 (20)
Median (IQR)	1 (1, 2)	5 (4, 7)	14 (11, 17)	10 (4, 19)	6 (2, 17)	34 (27, 45)	5 (2, 9)	15 (6, 30)
(10%, 90%)	(0, 3)	(3, 8)	(9, 21)	(2, 35)	(1, 29)	(21, 55)	(1, 15)	(1, 45)
Normalised brain volume								
Mean (SD)	1,541 (71)	1,481 (75)	1,450 (81)	1,496 (90)	1,505 (83)	1,411 (87)	1,467 (75)	1,396 (86)
Median (IQR)	1,541 (1,491, 1,592)	1,480 (1,430, 1,531)	1,447 (1,390, 1,512)	1,494 (1,437, 1,562)	1,507 (1,450, 1,566)	1,415 (1,347, 1,477)	1,462 (1,418, 1,512)	1,390 (1,349, 1,447)
(10%, 90%)	(1,447, 1,634)	(1,388, 1,583)	(1,350, 1,545)	(1,384, 1,610)	(1,400, 1,598)	(1,295, 1,508)	(1,374, 1,573)	(1,287, 1,513)
Number of T1 Gd lesions								
Mean (SD)	0 (0)	0 (0)	0 (0)	4 (7)	2 (5)	0 (0)	0 (0)	4 (13)
Median (IQR)	0 (0, 0)	0 (0, 0)	0 (0, 0)	2 (1, 5)	0 (0, 2)	0 (0, 0)	0 (0, 0)	0 (0, 2)
(10%, 90%)	(0, 0)	(0, 0)	(0, 0)	(1, 10)	(0, 6)	(0, 1)	(0, 0)	(0, 6)
RELAPSE	0 (0%)	0 (0%)	0 (0%)	0 (0%)	91 (100%)	0 (0%)	0 (0%)	0 (0%)

Demographic Summary: Baseline Meta states

```
baseline =follow%>%filter(DAY==1)
ff2<-baseline%>%dplyr::select(c("AGE", "SEX", "MSTYPE", "DURFS", "REL PST1Y", "MS"))%>%tbl_summary(by=MS, mis
```

```
label = 1.
type = li
statistic
```

```
ff2
```

```
## Table printed with `knitr::kable()`, not {gt}. Learn why at
## https://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include `message = FALSE` in code chunk header.
```

Variable	1, N = 1,059	2, N = 739	3, N = 91	4, N = 353
Age				
Mean (SD)	40 (9)	36 (9)	37 (9)	45 (8)
Median (IQR)	41 (34, 48)	36 (29, 43)	38 (31, 45)	46 (40, 51)
Gender				
F	657 (62%)	435 (59%)	66 (73%)	199 (56%)
M	402 (38%)	304 (41%)	25 (27%)	154 (44%)
MS Type				
PPMS	290 (27%)	176 (24%)	8 (8.8%)	236 (67%)
RMS	769 (73%)	563 (76%)	83 (91%)	117 (33%)
Duration from First Symptom				
Mean (SD)	6.4 (5.6)	5.9 (5.2)	6.1 (6.5)	8.7 (5.2)
Median (IQR)	5.0 (2.1, 8.9)	4.5 (1.8, 8.5)	3.9 (1.5, 8.3)	7.7 (4.9, 11.6)
Number of Relapses during past year				
Mean (SD)	1 (1)	1 (1)	1 (1)	0 (1)
Median (IQR)	1 (0, 1)	1 (0, 2)	1 (1, 2)	0 (0, 1)

Features Summary: Baseline Meta states

```
ff3<-baseline%>%dplyr::select(c("MSTYPE", "EDSS", "T25FWM", "HPT9M", "PASAT", "VOLT2", "NBV", "NUMGDT1", "RELAPSE"))
```

```
ff3
```

```
## Table printed with `knitr::kable()`, not {gt}. Learn why at
## https://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include `message = FALSE` in code chunk header.
```

Variable	1, N = 1,059	2, N = 739	3, N = 91	4, N = 353
MSTYPE				
PPMS	290 (27%)	176 (24%)	8 (8.8%)	236 (67%)
RMS	769 (73%)	563 (76%)	83 (91%)	117 (33%)
EDSS				
Mean (SD)	2.90 (1.21)	3.13 (1.51)	2.86 (1.34)	5.29 (1.14)
Median (IQR)	3.00 (2.00, 4.00)	3.00 (2.00, 4.00)	2.50 (2.00, 4.00)	5.50 (4.50, 6.00)
(10%, 90%)	(1.50, 4.50)	(1.50, 5.50)	(1.50, 4.50)	(3.50, 6.50)
Walking Time				
Mean (SD)	6 (3)	9 (10)	7 (8)	21 (27)
Median (IQR)	6 (5, 7)	6 (5, 9)	5 (4, 7)	12 (8, 22)
(10%, 90%)	(4, 9)	(4, 16)	(4, 9)	(6, 42)
Hand Coordination				

Variable	1, N = 1,059	2, N = 739	3, N = 91	4, N = 353
Mean (SD)	23 (5)	26 (9)	28 (30)	37 (28)
Median (IQR)	22 (20, 25)	24 (20, 29)	23 (20, 28)	31 (25, 39)
(10%, 90%)	(18, 29)	(18, 37)	(18, 32)	(22, 54)
PASAT				
Mean (SD)	44 (12)	41 (13)	42 (13)	36 (14)
Median (IQR)	47 (36, 54)	43 (32, 51)	47 (30, 52)	36 (26, 49)
(10%, 90%)	(25, 58)	(23, 57)	(24, 56)	(19, 55)
T2 lesion volume				
Mean (SD)	6 (6)	14 (15)	12 (14)	20 (19)
Median (IQR)	3 (1, 8)	10 (4, 19)	6 (2, 17)	14 (4, 32)
(10%, 90%)	(1, 14)	(2, 35)	(1, 29)	(1, 45)
Normalised brain volume				
Mean (SD)	1,501 (84)	1,496 (90)	1,505 (83)	1,438 (86)
Median (IQR)	1,504 (1,441, 1,560)	1,494 (1,437, 1,562)	1,507 (1,450, 1,566)	1,442 (1,381, 1,493)
(10%, 90%)	(1,391, 1,614)	(1,384, 1,610)	(1,400, 1,598)	(1,326, 1,535)
Number of T1 Gd lesions				
Mean (SD)	0 (0)	4 (7)	2 (5)	1 (4)
Median (IQR)	0 (0, 0)	2 (1, 5)	0 (0, 2)	0 (0, 0)
(10%, 90%)	(0, 0)	(1, 10)	(0, 6)	(0, 1)
RELAPSE	0 (0%)	0 (0%)	91 (100%)	0 (0%)

Including Plots

States density plots: baseline original features

```

p1<-ggplot(baseline, aes(y=as.factor(Z), x=EDSS, fill=as.factor(Z))) +
  geom_density_ridges(alpha=0.6, bandwidth=1) +
  scale_fill_viridis(discrete=TRUE) +
  scale_color_viridis(discrete=TRUE) +
  theme_bw() +
  theme(
    legend.position="none",
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  xlab("EDSS (total score)") +
  ylab("States")

p2<-ggplot(baseline, aes(y=as.factor(Z), x=T25FWM, fill=as.factor(Z))) +
  geom_density_ridges(alpha=0.6, bandwidth=4) +xlim(0,50)+
  scale_fill_viridis(discrete=TRUE) +
  scale_color_viridis(discrete=TRUE) +
  theme_bw() +
  theme(
    legend.position="none",
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  xlab("25-foot times walking test (s)") +

```

```

ylab("")

p3<-ggplot(baseline, aes(y=as.factor(Z), x=HPT9M, fill=as.factor(Z))) +
  geom_density_ridges(alpha=0.6, bandwidth=4) +xlim(0,100)+
  scale_fill_viridis(discrete=TRUE) +
  scale_color_viridis(discrete=TRUE) +
  theme_bw() +
  theme(
    legend.position="none",
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  xlab("Hand Coordination (s)") +
  ylab("")

p4<-ggplot(baseline, aes(y=as.factor(Z), x=PASAT, fill=as.factor(Z))) +
  geom_density_ridges(alpha=0.6, bandwidth=4) +xlim(0,60)+
  scale_fill_viridis(discrete=TRUE) +
  scale_color_viridis(discrete=TRUE) +
  theme_bw() +
  theme(
    legend.position="none",
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  xlab("PASAT (correct answers out of 60)") +
  ylab("")

p5<-ggplot(baseline, aes(y=as.factor(Z), x=NBV/1000, fill=as.factor(Z))) +
  geom_density_ridges(alpha=0.6, bandwidth=.05) +
  scale_fill_viridis(discrete=TRUE) +
  scale_color_viridis(discrete=TRUE) +
  theme_bw() +
  theme(
    legend.position="none",
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  xlab("Normalised Brain Volume (L)") +
  ylab("States")

p6<-ggplot(baseline, aes(y=as.factor(Z), x=VOLT2, fill=as.factor(Z))) +
  geom_density_ridges(alpha=0.6, bandwidth=.3) +xlim(NA,5)+
  scale_fill_viridis(discrete=TRUE) +
  scale_color_viridis(discrete=TRUE) +
  theme_bw() +
  theme(
    legend.position="none",
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  xlab("T2 Lesion Volume") +
  ylab("")

```



```
p7<-ggplot(baseline, aes(y=as.factor(Z), x=NUMGDT1, group=as.factor(Z))) +
  geom_density_ridges(aes(fill = as.factor(Z)), stat = "binline", binwidth = .5, scale = 0.95) +xlim(-1
  scale_fill_viridis(discrete=TRUE) +
  scale_color_viridis(discrete=TRUE) +
  theme_bw() +
  theme(
    legend.position="none",
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  xlab("Number of T1 Gd lesions") +
  ylab("")
```

```
p8<-ggplot(baseline, aes(y=as.factor(Z), x=RELAPSE, group=as.factor(Z))) +
  geom_density_ridges(aes(fill = as.factor(Z)), stat = "binline", binwidth = .5, scale = 0.95) +xlim(-1
  scale_fill_viridis(discrete=TRUE) +
  scale_color_viridis(discrete=TRUE) +
  theme_bw() +
  theme(
    legend.position="none",
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  xlab("Relapse Occurence") +
  ylab("")
```

```
(p1|p2|p3|p4)/(p5|p6|p7|p8)
```

```
## Warning: Removed 43 rows containing non-finite outside the scale range
## (`stat_density_ridges()`).
```

```
## Warning: Removed 10 rows containing non-finite outside the scale range
## (`stat_density_ridges()`).
```

```
## Warning: Removed 51 rows containing non-finite outside the scale range
## (`stat_density_ridges()`).
```

```
## Warning: Removed 1230 rows containing non-finite outside the scale range
## (`stat_density_ridges()`).
```

```
## Warning: Removed 82 rows containing non-finite outside the scale range
## (`stat_binline()`).
```

