

SUPPLEMENTARY MATERIAL

Dual Trajectories of Serum Brain-Derived Neurotrophic Factor and Cognitive Function in People Living with HIV

Henry U. Michael^{1,7,8} Antony M Rapulana^{2,5,6}, Theresa Smit², Njabulo Xulu², Sivapragashini Danaviah³, Suvira Ramlall⁴, Frasia Oosthuizen¹

1. Discipline of Pharmaceutical Sciences, School of Health Science, University of KwaZulu-Natal, Durban, South Africa
2. Africa Health Research Institute, Durban, South Africa
3. Faculty of Applied Science, Eduvos, Midrand, South Africa
4. Department of Psychiatry, University of KwaZulu-Natal, Durban, South Africa
5. School of Laboratory Medicine and Medical Science, University of Kwazulu-Natal, Durban, South Africa
6. UCL Centre for Clinical Microbiology, Division of Infection & Immunity, University College London, England
7. Centre for Outcomes Research & Evaluation, Research Institute of McGill University Health Centre (RI-MUHC), Montreal, Quebec, Canada
8. Division of Clinical and Translational Research, McGill University, Montreal, Quebec, Canada

Supplementary Material S1: Annotated Code for Dual Group-Based Trajectory Analysis

This write-up provides an explanation and breakdown of the annotated R code for conducting Group-Based Trajectory Analysis. The primary focus is on analyzing trajectories for NPZ-6 cognitive scores. Adaptations were used for BDNF and proBDNF biomarkers. Each section of the code is explained to highlight its purpose and ensure reproducibility.

Group-Based Trajectory Analysis for NPZ-6 #Adapt for BDNF and proBDNF

This script analyzes NPZ-6 cognitive scores using latent class linear mixed models.

Adaptations were also used for BDNF and proBDNF

Load necessary libraries

library(lcmm) # For latent class mixed models

library(dplyr) # For data manipulation

library(nnet) # For multinomial logistic regression

library(dunn.test) # For pairwise comparisons

library(readxl) # For reading external Excel files

1. Latent Class Linear Mixed Models Analysis -----

Linear model (1-class)

```
model1_linearNPZ <- hlme(fixed = Global_NPZ ~ Week,  
  random = ~ Week,  
  subject = 'Patid',  
  data = df)
```

```
summary(model1_linearNPZ)
```

```
# Add 2- and 3-class trajectory models
```

```
model2_linearNPZ <- hlme(fixed = Global_NPZ ~ Week,  
  random = ~ Week,  
  mixture = ~ Week,  
  subject = 'Patid',  
  ng = 2,  
  data = df,  
  B = random(model1_linearNPZ))
```

```
summary(model2_linearNPZ)
```

```
model3_linearNPZ <- hlme(fixed = Global_NPZ ~ Week,  
  random = ~ Week,  
  mixture = ~ Week,  
  subject = 'Patid',  
  ng = 3,  
  data = df,  
  B = random(model1_linearNPZ))
```

```
summary(model3_linearNPZ)
```

```
model4_linearBDNF <- hlme(fixed = log10BDNF ~ Week, random = ~ Week,
```

```
subject = 'Patid', ng = 4,  
mixture = ~ Week,  
data = df,  
B = random(model1_linearBDNF))  
summary(model4_linearBDNF)
```

Quadratic and cubic models #adapt for number of groups # and higher order-order polynomial relationships

Testing higher-order polynomial relationships for improved fit

```
model1_quadNPZ <- hlme(fixed = Global_NPZ ~ poly(Week, 2),  
random = ~ poly(Week, 2),  
subject = 'Patid',  
data = df)  
summary(model1_quadNPZ)
```

```
model3_cubicNPZ <- hlme(fixed = Global_NPZ ~ poly(Week, 3),  
random = ~ poly(Week, 3),  
mixture = ~ poly(Week, 3),  
subject = 'Patid',  
ng = 3,  
data = df,  
B = model1_cubicNPZ)  
summary(model3_cubicNPZ)
```

Model comparison using AIC and BIC to select the best-fitting model

```
summarytable(model1_linearBDNF, model1_quadBDNF, model1_cubicBDNF,  
             model2_linearBDNF, model2_quadBDNF, model2_cubicBDNF,  
             model3_linearBDNF, model3_quadBDNF, model3_cubicBDNF,  
             model4_linearBDNF, model4_quadBDNF, model4_cubicBDNF,  
             model5_linearBDNF, model5_quadBDNF, model5_cubicBDNF,  
             which = c("G", "loglik", "conv", "npm", "AIC", "BIC", "SABIC",  
                       "entropy", "ICL", "%class"))
```

2. Visualization of Best Model -----

Predicted trajectories for the best model (model3_cubicNPZ)

```
predIC1 <- predictY(model3_cubicNPZ, data_predNPZ, var.time = "Week", draws = TRUE)
```

Plot predicted trajectories

```
plot(predIC1, col = c("red", "navy", "darkgreen"), lty = 1, lwd = 3,  
     ylab = "Global_NPZ", main = "Predicted Trajectories for NPZ-6",  
     ylim = c(-5, 5), shades = TRUE, legend.loc = "topleft")
```

Add custom x-axis labels and legend

```
axis(1, at = c(0, 24, 48, 96), labels = c("0", "24", "48", "96"))  
legend("topleft", legend = c("Low baseline-slow improvement", "Gradual Improvement",  
                             "Late Surge"), col = c("red", "navy", "darkgreen"),  
      lty = 1, lwd = 3, bty = "n")
```

3. Posterior Probabilities and Odds of Classification -----

Compute posterior probabilities for class membership

```
posterior_probabilities <- postprob(model3_cubicNPZ)
```

Plot posterior probabilities

```
plot(model3_cubicNPZ, which = "postprob")
```

Mean posterior probabilities for each class

```
mean_posterior_probabilities <- c(class1 = 0.8820, class2 = 0.9094, class3 = 0.9492)
```

Odds of correct classification for each class

```
odds_of_correct_classification <- mean_posterior_probabilities / (1 - mean_posterior_probabilities)
```

```
print("Odds of Correct Classification:")
```

```
print(odds_of_correct_classification)
```

#identifying patient IDs

```
GBTM_NPZ <- model3_cubicNPZ$pprob[,1:2]
```

4. Statistical Analysis of Trajectory Classes -----

Kruskal-Wallis Test: E.g. Compare baseline age across trajectory classes

```
kruskal_test_age <- kruskal.test(Baseline_age ~ NPZ_trajclass, data = df)
```

```
print(kruskal_test_age)
```

Dunn's Test: Post-hoc analysis for pairwise comparisons

```
dunn_result_age <- dunn.test(df$Baseline_age, df$NPZ_trajclass, method = "holm")
```

```
print(dunn_result_age)
```

5. Multinomial Logistic Regression -----

*Predictors of trajectory class membership*

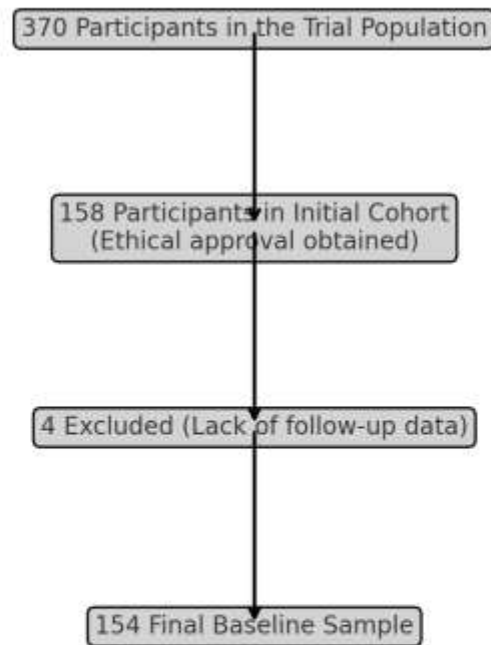
```
multinom_model <- multinom(NPZ_trajclass ~ Baseline_age + Sex + Edu_yrs +  
    CD4_CD8_ratio + log10ProBDNF + Global_NPZ,  
    data = NPZtraj)  
  
summary(multinom_model)  
  
# Odds ratios from multinomial logistic regression  
exp(coef(multinom_model))
```

6. Conditional Probabilities -----

*Compute joint and conditional probabilities for trajectory group memberships*

```
contingency_table <- table(df_class$ProBDNFclass, df_class$class.y)  
conditional_prob_Global_NPZ_given_logProBDNF <- prop.table(contingency_table, 1)  
conditional_prob_logProBDNF_given_Global_NPZ <- prop.table(contingency_table, 2)  
print("Conditional Probabilities of Global NPZ given logProBDNF:")  
print(conditional_prob_Global_NPZ_given_logProBDNF)  
print("Conditional Probabilities of logProBDNF given Global NPZ:")  
print(conditional_prob_logProBDNF_given_Global_NPZ)  
  
# Joint probabilities  
joint_probabilities <- prop.table(contingency_table)  
print("Joint Probabilities of Trajectory Group Memberships:")  
print(joint_probabilities)
```

S2 Flow diagram summarizing the study participant inclusion and exclusion process.



S3 Determination of number and shape of mBDNF Trajectory Groups and Model diagnostics

Number of Groups	Shape	BIC, Participants (N = 155)	BIC, Observations (N = 562)	Entropy	Class 1 %	Class 2 %	Class 3 %	Class 4 %	Minimum average posterior probability	Minimum Odds of correct classification
1	Linear	-327.1463	-319.41	1.0000000	100	-	-	-	-	-
1	Quadratic	-341.2772	-328.41	1.0000000	100	-	-	-	-	-
1	Cubic	-337.7450	-318.43	1.0000000	100	-	-	-	-	-
2	Linear	-312.0160	-300.42	0.0005258	48.39	51.61	-	-	0.51	1.04
2	Quadratic	-321.1034	-303.08	0.0015496	45.16	54.84	-	-	0.51	1.04
2	Cubic	-331.0755	-305.31	0.7865736	16.13	83.87	-	-	0.91	9.74
3	Linear	-315.7490	-298.82	0.8812609	95.48	2.58	1.94	-	0.74	9.74
3	Quadratic	-355.2186	-332.03	0.9798379	1.29	98.06	0.65	-	0.99	99
3	Cubic	-348.0031	-315.79	0.9700322	1.29	94.84	3.87	-	0.86	6.4
4	Linear	-299.1428	-279.83	0.2540520	0	97.42	2.58	0	NA	NA
4	Quadratic	-340.6700	-312.33	0.9410512	1.29	0.65	95.48	2.58	0.73	2.77
4	Cubic	-317.1501	-278.51	0.7225131	3.87	12.26	37.42	46.45	0.82	4.74

BIC - Bayesian Information Criterion

Values highlighted in red shows values that failed to meet our apriori criteria

Value highlighted in **bold** and green shows our chosen model as it fits all criteria

S4 Multivariable logistic regression of mBDNF Trajectory Group Membership

Variable	Estimate (log odds)	Std. Error	p-value
Age	-0.11	0.04	0.01*
CD8	-0.001	0.0004	0.02*
log10 BDNF log 10	5.23	1.74	0.003**
ProBDNF	-0.31	0.41	0.45
log RNA	-0.01	0.36	0.97
Education	-0.02	0.1	0.86
Sex (Male)	0.57	0.65	0.38
Site (Harare)	-0.38	0.62	0.55

BDNF, Brain-Derived Neurotrophic Factor; proBDNF, precursor Brain-Derived Neurotrophic Factor; RNA, Viral Load

S5 Determination of the number and shape of ProBDNF Trajectory Groups and Model diagnostics

Number of Groups	Shape	BIC, Participants (N = 154)	BIC, Observations (N = 486)	Entropy	Class 1 %	Class 2 %	Class 3 %	Class 4 %	Minimum average posterior probability	Minimum Odds of correct classification
1	Linear	1066.52	1073.42	1.0000000	100	-	-	-	-	-
1	Quadratic	1068.8	1080.3	1.0000000	100	-	-	-	-	-
1	Cubic	1089.76	1106.99	1.0000000	100	-	-	-	-	-
2	Linear	10078.78	1089.12	0.4521305	11.04	88.96	-	-	0.69	2.25
2	Quadratic	1083.99	1100.07	0.5074554	10.4	89.61	-	-	0.76	3.24
2	Cubic	1105.42	1128.4	0.5395701	14.29	85.71	-	-	0.86	5.92
3	Linear	1086.65	1100.44	0.4160715	35.71	38.31	25.97	-	0.64	1.77
3	Quadratic	1094.4	1115.09	0.4955077	12.34	45.45	42.21	-	0.72	2.6
3	Cubic	1120.19	1148.92	0.5104648	18.18	52.6	29.22	-	0.65	1.87
4	Linear	1100.25	1117.49	0.5505861	36.36	0.65	37.01	25.97	0.67	2.03
4	Quadratic	1105.49	1130.77	0.5027254	19.48	37.66	16.23	26.62	0.63	1.7
4	Cubic	1125.52	1161.99	0.6252928	16.88	9.09	51.95	22.97	0.66	1.93

BIC - Bayesian Information Criterion

Values highlighted in red shows values that failed to meet our apriori criteria

Value highlighted in bold and green shows our chosen model as it fits all criteria

S6 Multivariable logistic regression of proBDNF Trajectory Group Membership

Variable	Estimate (log odds)	Std. Error	p-value
Age	-0.07	0.04	0.08
CD8	-0.0004	0.0005	0.43
log10 BDNF	0.07	1.41	0.96
log 10			
ProBDNF	0.24	0.38	0.53
log RNA	0.36	0.34	0.31
Education	-0.09	0.11	0.42
Sex (Male)	0.21	0.62	0.74
Site (Harare)	1.91	0.7	0.007**

BDNF, Brain-Derived Neurotrophic Factor; proBDNF, precursor Brain-Derived Neurotrophic Factor; RNA, Viral Load

S7 Determination of the number and shape of NPZ-6 Trajectory Groups and Model diagnostics

Number of Groups	Shape	BIC, Participants (N = 154)	BIC, Observations (N =486)	Entropy	Class 1 %	Class 2 %	Class 3 %	Class 4 %	Minimum average posterior probability	Minimum Odds of correct classification
1	Linear	1440.86	1448.91	1.0000000	100	-	-	-	-	-
1	Quadratic	1418.9	1432.31	1.0000000	100	-	-	-	-	-
1	Cubic	1431.17	1451.3	1.0000000	100	-	-	-	-	-
2	Linear	1456	1431.51	0.4521305	38.06	61.94	-	-	0.53	1.14
2	Quadratic	1439.07	1457.85	0.5074554	92.26	7.74	-	-	0.51	1.06
2	Cubic	1441.27	1468.1	0.5395701	75.48	24.52	-	-	0.88	7.55
3	Linear	1424.23	1440.32	0.4160715	93.55	0.65	5.81	-	0.84	5.3
3	Quadratic	1430.62	1454.77	0.4955077	27.1	63.87	9.03	-	0.84	5.44
3	Cubic	1447.68	1481.23	0.5104648	24.52	67.74	7.74	-	0.88	7.47
4	Linear	1430.01	1450.14	0.5505861	26.45	70.32	0.65	2.58	0.64	1.8
4	Quadratic	1450.79	1480.31	0.5027254	30.32	59.35	0	10.32	0	0
4	Cubic	1432.16	1472.42	0.6252928	1.29	54.84	32.9	10.97	0.91	10.09

BIC - Bayesian Information Criterion

Values highlighted in red shows values that failed to meet our apriori criteria

Value highlighted in **bold** and green shows our chosen model as it fits all criteria

S8 Identifying predictors of NPZ-trajectory group membership using Multinomial logistic regression

Variable	Estimate	Std. Error	z value	Pr(> z)	Significance
(Intercept):1	10.57843	3.98511	2.654	0.00794	**
(Intercept):2	3.5222	3.07114			
Baseline_age:1	-0.21448	0.07404	-2.897	0.00377	**
Baseline_age:2	-0.01754	0.04885	-0.359	0.71948	
SexM:1	2.26779	1.2445	1.822	0.06842	.
SexM:2	1.50866	1.09348	1.38	0.16768	
Edu_yrs:1	-0.27413	0.17158	-1.598	0.11012	
Edu_yrs:2	-0.01614	0.14401	-0.112	0.91076	
CD4_CD8_ratio:1	-2.55241	3.32758	-0.767	0.44305	
CD4_CD8_ratio:2	-1.00751	2.7069	-0.372	0.70974	
log10ProBDNF:1	0.06197	0.58094	0.107	0.91505	
log10ProBDNF:2	-0.31349	0.48024	-0.653	0.5139	
Global_NPZ:1	-1.8881	0.60828	-3.104	0.00191	**
Global_NPZ:2	-0.09398	0.48161	-0.195	0.84528	

BDNF, Brain-Derived Neurotrophic Factor; proBDNF, precursor Brain-Derived Neurotrophic Factor; RNA, Viral Load, Global Composite score, Global_NPZ

S9 Interrelationships Between Serum mBDNF and proBDNF Trajectories

The linkage between Serum mBDNF and ProBDNF trajectory groups presented by conditional and joint probabilities from the group-based dual trajectory model.

3a Probability of ProBDNF group conditional on serum mBDNF group			
	ProBDNF trajectory group		
Serum mBDNF trajectory group	Gradual Decline	Gradual Increase	Total
Peak with gradual decline	0.12	0.88	1
Stable Ascent	0.15	0.85	1
3b Probability of Serum mBDNF group conditional on ProBDNF trajectory group			
	mBDNF trajectory group		
Serum proBDNF trajectory group	Peak with gradual decline	Stable Ascent	
Gradual Decline	0.14	0.86	1
Gradual Increase	0.18	0.82	1
3c Joint probability of Serum mBDNF and ProBDNF trajectory groups			
	ProBDNF trajectory groups		
Serum mBDNF trajectory group	Gradual Decline	Gradual Increase	
Peak with gradual decline	0.02	0.11	
Stable Ascent	0.13	0.56	

Notes: **S7a**, Conditional probability of an individual being in a specific serum proBDNF trajectory group given their serum mBDNF trajectory group. **S7b**, Conditional probability of an individual being in a specific serum mBDNF trajectory group given their serum proBDNF trajectory group. **S7c**, Joint probabilities, combining the two aspects (serum mBDNF trajectory and serum proBDNF trajectory) to indicate the probability of both occurring together. Abbreviations: mBDNF, mature Brain-derived Neurotrophic Factor; ProBDNF, Precursor Brain-derived Neurotrophic Factor; Values in **bold** represent high conditional and joint probabilities.

