LOGBOOK FOR "PTSD SYMPTOM TRAJECTORIES IN COMMUNITY VOLUNTEERS"

Rens van de Schoot
20-01-2015

CONTENTS

1. SPSS DATA

The SPSS data file was received on 14-01-2013, see "Data FOR JAMA ARTICLE_14.01.2013.sav" with descriptions of the variable available in "JAMA_excel explanations data.xlsx". For all variables missing data was recoded into -999 and the spss file was saved as tab-delimneted file, see "Data FOR JAMA ARTICLE_14.01.2013.dat".

2. LGMM TOTAL SAMPLE

The total sample has been used to estimate latent growth trajectories with 1-6 latent groups:

```
MODEL: !here you specify the model to be estimated

%overall%
Intercept Slope |
IES_I@0
IES_II@6
IES_III@12;

intercept with slope @0;
```

See the files "LGMM_2 nr of classes is 1-3".

Nr. Of classes	AIC	BIC	BLRT	VLMR	Entropy	Sample size per class
1	8623	8656	-	-	-	509
2	8609	8656	P<.001	0.0484	0.583	452/57
3[a]	8605	8660	0.1364	0.1700	0.703	449/50/10
4[a]	8600	8668	0.0619	0.5014	0.567	224/223/48/14

a= WARNING: THE LATENT VARIABLE COVARIANCE MATRIX (PSI) IN CLASS 1
IS NOT POSITIVE DEFINITE. THIS COULD INDICATE A NEGATIVE VARIANCE/
RESIDUAL VARIANCE FOR A LATENT VARIABLE, A CORRELATION GREATER OR EQUAL
TO ONE BETWEEN TWO LATENT VARIABLES, OR A LINEAR DEPENDENCY AMONG MORE
THAN TWO LATENT VARIABLES. CHECK THE TECH4 OUTPUT FOR MORE INFORMATION.
PROBLEM INVOLVING VARIABLE SLOPE.

Conclusion: the three class model does not work out, but the two class model is not that good according to the entropy.

Results not reported in main text.

3. LGMM TOTAL SAMPLE WITH QUADRATIC TREND

A quadratic trend has been added to the model, but to simplify the model and to avoid computational issues the variance around the quadratic trend has been fixed to zero:

```
MODEL: !here you specify the model to be estimated

%overall%
Intercept Slope Q|
IES_I@0
IES_II@6
IES_III@12;

q@0;
```

See the files “lgmm_2 nr of classes is 1-3 q”

Nr. Of classes	Q term significant?	AIC	BIC	BLRT	VLMR	Entropy	Sample size per class
1	0.906	8625	8663	-	-	-	509
2	0.375/0.205	8612	8667	0.0128	0.0756	0.544	438/71
3	0.003/0.000/0.000	8594	8662	<.001	0.0006	0.682	428/68/13
4[a]	0.000/0.857/0.209/0.934	8591	8676	0.2000	0.3358	0.635	240/222/31/16

a = random starts were increased because the best log likelihood was not replicated

Conclusion: q does not fit the data better

Results not reported in main text.

4. LGMM PER GROUP QUADRATIC

The LGMM analyses were repeated but now per group with a quadratic trend but to simplify the model and to avoid computational issues the variance around the quadratic trend has been fixed to zero:

```
MODEL: !here you specify the model to be estimated

%overall%
Intercept Slope q|
IES_I@0
IES_II@6
IES_III@12;

q@0;
```

See the files “lgmm nr of classes is 1-4 knownclass”.

Nr. Of classes	AIC	BIC	BLRT[a]	VLMR[b]	Entropy	Sample size per class
1	8459	8512	-	-	-	220 229
2	8450	8532			0.804	119/21 205/24
3	8447	8558			0.702	106/94/20 11/94/24
4[c]	8443	8582			0.741	102/89/24/5 11/92/23/3

a = *** WARNING in OUTPUT command
TECH11 option is not available for TYPE=MIXTURE with the KNOWNCLASS option.
Request for TECH11 is ignored.

b = *** WARNING in OUTPUT command
TECH14 option is not available for TYPE=MIXTURE with more than one
categorical latent variable. Request for TECH14 is ignored.

c = random starts were increased because the best log likelihood was not replicated, but still warnings:
WARNING: THE RESIDUAL COVARIANCE MATRIX (THETA) IN CLASS 1 IS NOT
POSITIVE DEFINITE. THIS COULD INDICATE A NEGATIVE VARIANCE/RESIDUAL
VARIANCE FOR AN OBSERVED VARIABLE, A CORRELATION GREATER OR EQUAL TO ONE
BETWEEN TWO OBSERVED VARIABLES, OR A LINEAR DEPENDENCY AMONG MORE THAN TWO
OBSERVED VARIABLES. CHECK THE RESULTS SECTION FOR MORE INFORMATION.
PROBLEM INVOLVING VARIABLE IES_III.

Data set contains unknown or missing values for variable YEARVOLUNT_CAT.
This variable is used to determine the KNOWNCLASS specification.
Number of such cases: 65

N for analyses: n=449

5. LGMM PER GROUP LINEAR

The LGMM analyses were repeated but now per group with a linear trend but to simplify the model and to avoid computational issues the variance around the quadratic trend has been fixed to zero. See "LGMM nr of classes is 1-4 KNOWNCLASS LINEAR".

	Akaike (AIC)	Bayesian (BIC)	Entropy	Sample sizes Non-core volunteers	Sample sizes Core volunteers
1 class	8455	8500	1	220	229
2 classes	8442	8508	0.806	199/21	205/24
2 classes[b]	8439	8497	0.810	202/18	206/23
3 classes[a]	8437	8523	0.707	107/94/19	113/93/23
3 classes[b]	8438	8516	0.832	202/15/3	203/22/4
4 classes[a]	8436	8543	0.739	105/92/15/8	111/92/22/4
4 classes[b]	8438	8537	0.712	112/82/21/5	119/83/22/5

[a] Non-positive definite matrix because of negative slope variance
[b] slope variance fixed to zero

*** WARNING in OUTPUT command
TECH11 option is not available for TYPE=MIXTURE with the KNOWNCLASS option.
Request for TECH11 is ignored.

*** WARNING in OUTPUT command
TECH14 option is not available for TYPE=MIXTURE with more than one
categorical latent variable. Request for TECH14 is ignored.

a = WARNING: THE LATENT VARIABLE COVARIANCE MATRIX (PSI) IN CLASS 1
IS NOT POSITIVE DEFINITE. THIS COULD INDICATE A NEGATIVE VARIANCE/
RESIDUAL VARIANCE FOR A LATENT VARIABLE, A CORRELATION GREATER OR EQUAL
TO ONE BETWEEN TWO LATENT VARIABLES, OR A LINEAR DEPENDENCY AMONG MORE
THAN TWO LATENT VARIABLES. CHECK THE TECH4 OUTPUT FOR MORE INFORMATION.
PROBLEM INVOLVING VARIABLE SLOPE.

Data set contains unknown or missing values for variable YEARVOLUNT_CAT.
This variable is used to determine the KNOWNCLASS specification.
Number of such cases: 65

N for analyses: n=449

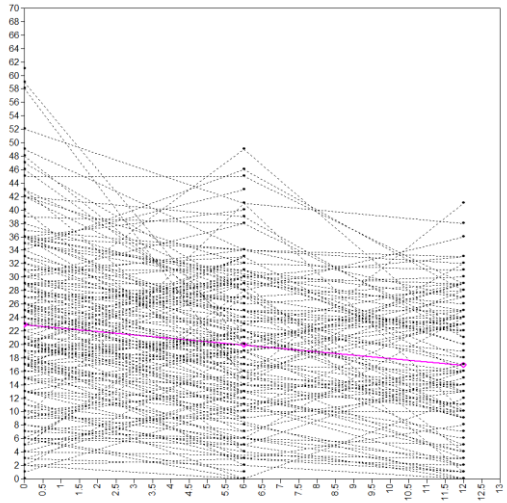
6. Output Final Model

	Resilient (n=206/202)				Chronic (n=23/18)			
	M	SE	95% CI	p	M	SE	95% CI	p
Core volunteers								
Intercept	22.951	0.961	21.067 - 24.835	<.001	27.662	2.857	22.062 - 33.262	<.001
Slope	-0.505	0.085	-0.673 - -0.338	<.001	1.403	0.339	0.738 - 2.068	<.001
Non-core volunteers								
Intercept	27.901	1.995	23.991 - 31.810	<.001	31.050	9.248	12.925 - 49.176	.001
Slope	-0.486	0.110	-0.702 - -0.270	<.001	0.960	0.634	-0.284 - 2.203	.130

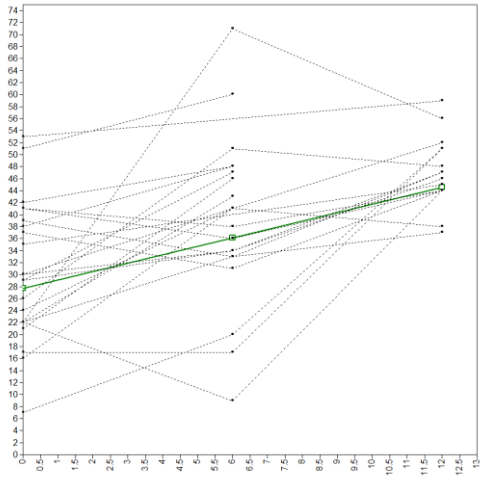
Effect sizes calculated based on equation 6 this paper: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2712654/ => mean difference between intercept parameters divided by the variance of the intercept parameter

Non core chronic versus resilient => (31.050-27.901)/sqrt(70.981) = .37
core chronic versus resilient => (27.662-22.951)/sqrt(70.981) = .56

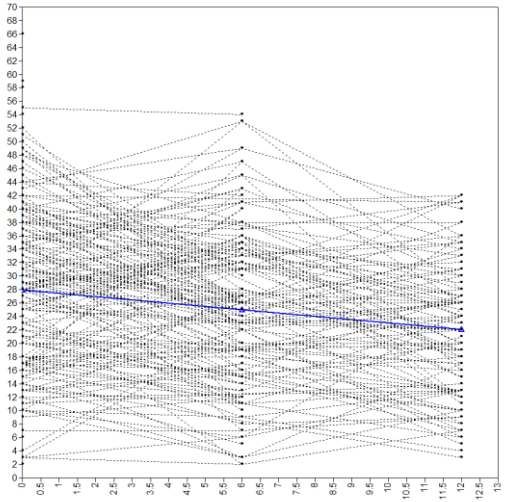
non-core chronic versus core chronic => (31.050-27.662)/sqrt(70.981) = .40
non-core resilient versus core resilient => (27.901- 22.901)/sqrt(70.981) = .59



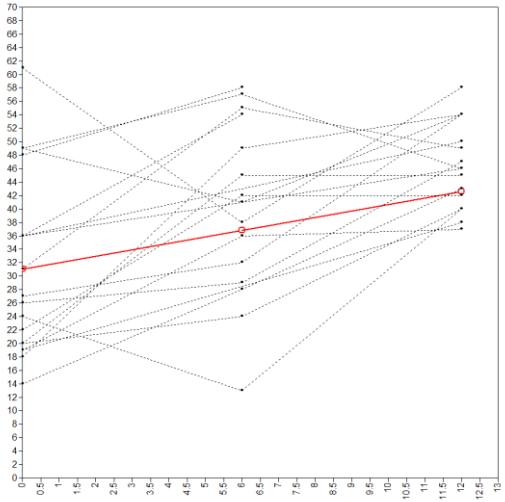
Core Resilient volunteers



Core Chronic volunteers



Non-core Resilient volunteers



Non-core chronic volunteers

Figure X. Estimated mean trajectories and observed individual trajectories.

Average Latent Class Probabilities for Most Likely Latent Class Pattern (Row)
by Latent Class Pattern (Column)

Latent Class Variable Patterns

Latent Class Pattern No.	G Class	C Class
1	1	1
2	1	2
3	2	1
4	2	2

	1	2	3	4
1	0.772	0.228	0.000	0.000
2	0.104	0.896	0.000	0.000
3	0.000	0.000	0.806	0.194
4	0.000	0.000	0.079	0.921

CLASSIFICATION OF INDIVIDUALS BASED ON THEIR MOST LIKELY LATENT CLASS PATTERN

Class Counts and Proportions

Latent Class Pattern			
1	1	18	0.04009
1	2	202	0.44989
2	1	23	0.05122
2	2	206	0.45880

CLASSIFICATION OF INDIVIDUALS BASED ON THEIR MOST LIKELY LATENT CLASS MEMBERSHIP
FOR EACH LATENT CLASS VARIABLE

Latent Class Variable	Class		
G	1	220	0.48998
	2	229	0.51002
C	1	41	0.09131
	2	408	0.90869

7. MODEL TEST BETWEEN 4 CLASSES

To test differences between intercepts and slopes, the following syntax has been used:

```
%g#1.c#1% !non-core + chronic
  [intercept](i1);
  [slope](s1);

%g#2.c#1% !core chronic
  [intercept](i2);
  [slope](s2);

%g#1.c#2% !non-core resilient
  [intercept](i3);
  [slope](s3);

%g#2.c#2% !core resilient
  [intercept](i4);
  [slope](s4);

MODEL TEST:
  i1=i2;
```

Results:

i1 = i2 -> file “modelt test 1.inp”-> p= 0.7032
i1 = i3 -> file “modelt test 2.inp”-> p= 0.7750
i1 = i4 -> file “modelt test 3.inp”-> p= 0.3966
i2 = i3 -> file “modelt test 4.inp”-> p= 0.9512
i2 = i4 -> file “modelt test 5.inp”-> p= 0.1506
i3 = i4 -> file “modelt test 6.inp”-> p= 0.0120

s1 = s2 -> file “modelt test 7.inp”-> p= 0.4452
s1 = s3 -> file “modelt test 8.inp”-> p= 0.0331
s1 = s4 -> file “modelt test 9.inp”-> p= 0.0186
s2 = s3 -> file “modelt test 10.inp”-> p<.001
s2 = s4 -> file “modelt test 11.inp”-> p<.001
s3 = s4 -> file “modelt test 12.inp”-> p= 0.8894

8. PREDICTING CLASS MEMBERSHIP

Since some predicting variables were not included in the file I received (see chapter 1), therefore I requested a new data set with the following variables:

TESS_OCC_ResourceLoss
educ_level_I
VAR00001
selfeff_I
Pre_emotional_I
t_PSP_I
t_evacuation_I

This data was sent by Sirry on 09-01-2015 in the file titled “DATA RENS.sav” and the variables were merged (on ID) in the file titled “Data FOR JAMA ARTICLE_14.01.2013 - including new variables.sav”. Missing data was recoded and the file was saved as tabdelimneted under the name “Data FOR JAMA ARTICLE_09-01-2014.dat”.

Running the model with the knownclass option resulted in the following error:

```
*** ERROR in VARIABLE command
  Auxiliary variables with E, R, R3STEP, DU3STEP, DE3STEP, DCATEGORICAL, or DCONTINUOUS
  are not available with TYPE=MIXTURE with more than one categorical latent variable.
```

Therefore we ran the model separately for core and non-core volunteers using the syntax:

```
USEOBSERVATION ARE (Yearvolunt_cat EQ 1);
```

Results non-core volunteers:

See the Mplus files “Three step method - Yearvolunt_cat EQ 1 .inp”

```
 WARNING:  LISTWISE DELETION IS APPLIED TO THE AUXILIARY VARIABLES IN THE
   ANALYSIS.  TO AVOID LISTWISE DELETION, DATA IMPUTATION CAN BE USED
   FOR THE AUXILIARY VARIABLES FOLLOWED BY ANALYSIS WITH TYPE=IMPUTATION.
   NUMBER OF DELETED OBSERVATIONS:  103
   NUMBER OF OBSERVATIONS USED:  117
```

⇨ Cause is SA_SUM_I

				Two-Tailed
	Estimate	S.E.	Est./S.E.	P-Value
C#1 ON				
AGE_I_CO	-0.159	0.113	-1.397	0.162
GENDER_I	0.327	4.845	0.067	0.946
SA_SUM_I	0.133	0.155	0.853	0.394
SELFEFF_	3.568	5.343	0.668	0.504
PRE_EMOT	**25.027**	**0.000**	**999.000**	**0.000**
T_PSP_I	-1.264	2.220	-0.570	0.569
T_EVACUA	3.025	4.553	0.664	0.507

Results without SA_SUM_I:
See the file "Three step method - Yearvolunt_cat EQ 1 - without SA_sum.inp"

```
WARNING:  LISTWISE DELETION IS APPLIED TO THE AUXILIARY VARIABLES IN THE
   ANALYSIS.  TO AVOID LISTWISE DELETION, DATA IMPUTATION CAN BE USED
   FOR THE AUXILIARY VARIABLES FOLLOWED BY ANALYSIS WITH TYPE=IMPUTATION.
   NUMBER OF DELETED OBSERVATIONS:  23
   NUMBER OF OBSERVATIONS USED:  197
```

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
C#1 ON				
AGE_I_CO	-0.030	0.062	-0.484	0.628
GENDER_I	0.191	0.877	0.218	0.827
SELFEFF_	0.785	0.630	1.247	0.212
PRE_EMOT	0.892	1.065	0.838	0.402
T_PSP_I	0.452	1.059	0.426	0.670
T_EVACUA	0.099	1.222	0.081	0.936

Results Core volunteers:
See the file "Three step method - Yearvolunt_cat EQ 2 .inp"

```
WARNING:  LISTWISE DELETION IS APPLIED TO THE AUXILIARY VARIABLES IN THE
   ANALYSIS.  TO AVOID LISTWISE DELETION, DATA IMPUTATION CAN BE USED
   FOR THE AUXILIARY VARIABLES FOLLOWED BY ANALYSIS WITH TYPE=IMPUTATION.
   NUMBER OF DELETED OBSERVATIONS:  128
   NUMBER OF OBSERVATIONS USED:  101
```

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
C#1 ON				
AGE_I_CO	0.111	0.084	1.324	0.185
GENDER_I	-1.223	1.419	-0.862	0.389
SA_SUM_I	**-0.296**	**0.152**	**-1.953**	**0.051**
SELFEFF_	-0.934	0.915	-1.021	0.307
PRE_EMOT	2.391	1.886	1.267	0.205
T_PSP_I	**-1959.592**	**1.615**	**-1213.558**	**0.000**
T_EVACUA	**1955.730**	**0.000**	**999.000**	**0.000**

Results without SA_SUM_I:
See the file "Three step method - Yearvolunt_cat EQ 2 - without SA_sum.inp"

```
WARNING:  LISTWISE DELETION IS APPLIED TO THE AUXILIARY VARIABLES IN THE
   ANALYSIS.  TO AVOID LISTWISE DELETION, DATA IMPUTATION CAN BE USED
   FOR THE AUXILIARY VARIABLES FOLLOWED BY ANALYSIS WITH TYPE=IMPUTATION.
   NUMBER OF DELETED OBSERVATIONS:  24
   NUMBER OF OBSERVATIONS USED:  205
```

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
C#1 ON				
AGE_I_CO	-0.028	0.044	-0.644	0.519
GENDER_I	-0.535	0.986	-0.542	0.588
SELFEFF_	-1.090	1.310	-0.832	0.405
PRE_EMOT	**-1.576**	**0.722**	**-2.183**	**0.029**
T_PSP_I	**-1.964**	**0.836**	**-2.350**	**0.019**
T_EVACUA	1.096	0.770	1.423	0.155

9. SAVING MEMBERSHIP

Using the SAVEDATA command the most likely membership classification was saved and transposed to an SPSS data set and merged with the original data. See the files in this subfolder.

Saved membership: see file save membership.inp

Data imported in SPSS, see file membership 2.sav

Kept only the variables ID, c, and MLCJOINT

Merged with the file “Data FOR JAMA ARTICLE_14.01.2013 - including new variables.sav”, see file “Data FOR JAMA ARTICLE_14.01.2013 - including new variables - including membership.sav”

```
DATASET ACTIVATE DataSet4.
MATCH FILES /FILE=*
  /FILE='DataSet5'
  /BY ID.
EXECUTE.

RECODE Chronic MLCJOINT (MISSING=-999).
EXECUTE.
```

10. POST-HOC ANALYSES

Marit Sijbrandij added the membership variables to the data set “Total file merged.sav” and ran the analyses as described in the syntax file “Syntax covariates.sps”