

Supplementary Materials for
**Arginine vasopressin in cerebrospinal fluid is a marker of sociality in
nonhuman primates**

Karen J. Parker,* Joseph P. Garner, Ozge Oztan, Erna R. Tarara, Jiang Li,
Valentina Sclafani, Laura A. Del Rosso, Katie Chun, Sean W. Berquist,
Michael G. Chez, Sonia Partap, Antonio Y. Hardan, Elliott H. Sherr, John P. Capitanio

*Corresponding author. Email: kjparker@stanford.edu

Published 2 May 2018, *Sci. Transl. Med.* **10**, eaam9100 (2018)
DOI: 10.1126/scitranslmed.aam9100

This PDF file includes:

- Fig. S1. Receiver operating characteristic for the discriminant analysis in the monkey discovery cohort.
- Table S1. Raw data for fig. S1.
- Table S2. SAS code for the data and analysis shown in Fig. 1.
- Table S3. SAS code for the data and analysis shown in Fig. 2.
- Table S4. SAS code for the data and analysis shown in Fig. 3A.
- Table S5. Confirmation of the specificity and validity of the statistical winnowing strategy for the monkey discovery cohort.
- Table S6. SAS code for the data and analysis shown in Fig. 4.
- Table S7. SAS code for the data and analysis shown in Fig. 3B.
- Table S8. SAS code for the data and analysis shown in Fig. 5.

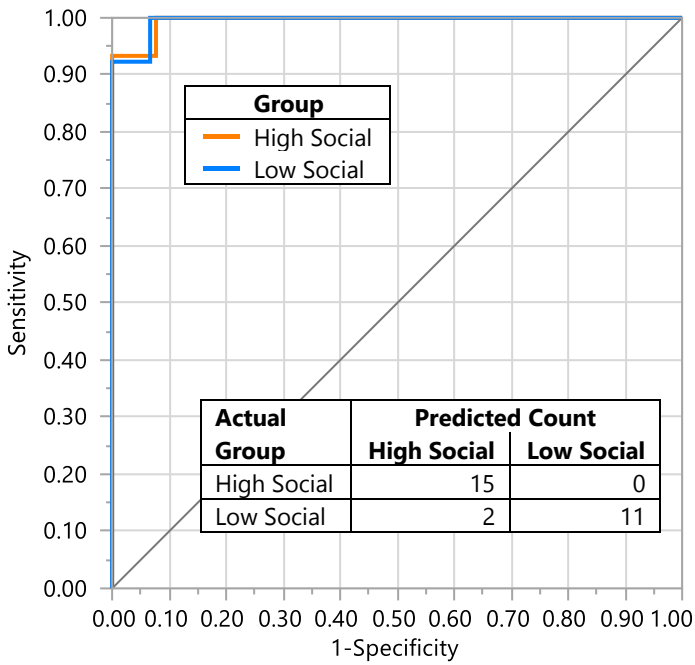


Figure S1. Receiver operating characteristic for the discriminant analysis in the monkey discovery cohort. The ROC plots the ability of the discriminant analysis to correctly classify individual monkeys into social groups (i.e., low-social vs. high-social) on the basis of the biological measures. Two monkey subjects were excluded due to missing biological data. Twenty-six out of 28 individuals (93%) were correctly classified as shown. Data were tested using LR ChiSq; $p < 0.0001$. Low-social monkeys are depicted in blue and high-social monkeys are depicted in orange.

Table S1. Raw data for fig. S1.

Group	CSF_OXT	CSF_AVP	Plasma_OXT	Plasma_AVP	OXTR_Delta_Ct	AVPR1A_Delta_Ct	log10(p-ERK/ERK)	log10(p-PTEN/PTEN)	log10(p-AKT1/AKT1)
High Social	46.761	6.22	12.801	8.318	-11.37030474	-11.78236453	-1.071071926	-0.085288548	-2.176871523
Low Social	46.013	5.942	4.884	15.351	-9.044528961	-9.591985703	-1.336504636	-0.190604563	-2.224176352
High Social	61.547	8.644	2.273	8.326	-7.62503624	-10.21472104	-1.381031155	-0.276778262	-2.227211819
High Social	50.857	7.124	13.09	7.304	-11.55533536	-12.47716904	-0.982959586	-0.165349554	-2.177361891
Low Social	86.68	3.152	8.233	18.145	-10.37744427	-12.06772804	-0.999604537	-0.299552308	-1.970982412
High Social	41.029	5.206	7.853	9.384	-9.57947731	-11.14218044	-1.36336478	-0.295737827	-2.330622102
Low Social	46.967	6.238	7.554	5.194	-11.28111966	-11.53748894	-1.27719587	-0.209324446	-2.016570889
Low Social	43.645	6.691	10.08	36.271	-11.48736699	-13.40877342	-1.588353108	-0.131656353	-2.113121894
Low Social	53.034	4.531	8.981	5.613	-10.99831486	-11.99593258	-1.449936085	-0.207848822	-2.056500073
Low Social	39.24	6.135	6.782	5.403	-10.18530973	-12.43996112	-1.244896767	-0.234022353	-1.948415132
Low Social	27.207	4.567	3.978	8.774	-11.4355348	-12.39903641	-1.509921076	-0.152687103	-2.150589719
Low Social	84.82	6.678	14.905	13.286	-10.58326467	-12.86216609	-1.720108738	-0.031438263	-1.979683372
High Social	46.939	5.79	1.082	7.669	-12.34000778	-12.88657888	-0.965661088	-0.22972812	-1.946674508
High Social	41.647	2.901	7.604	11.113	-10.87373352	-11.56472015	-1.397465756	-0.343990687	-2.14189503
High Social	61.653	5.692	2.942	6.405	-9.878153483	-10.24778175	-1.005154551	-0.225806599	-2.033086494
High Social	42.078	6.859	6.239	8.749	-11.56576157	-12.24991226	-1.426260817	-0.329361239	-2.130418198
Low Social	37.623	6.036	10.944	9.687	-10.21809069	-12.37737592	-1.217354598	-0.125400566	-2.118912255
High Social	59.596	7.43	2.101	5.661	-12.25979614	-12.6275177	-1.150048643	-0.181732505	-1.958504188
Low Social	35.517	6.354	11.663	7.123	-11.07358551	-11.01345952	-1.128011746	-0.300877132	-2.116930029
Low Social	40.439	8.229	3.542	8.486	-10.2171332	-10.65294393	-1.117957225	-0.219141039	-2.158332216
Low Social	51.123	5.996	6.537	7.825	-10.07591629	-10.4040858	-1.048958974	-0.146940046	-2.203823567
High Social	39.297	9.484	11.59	7.847	-11.20992088	-12.51468468	-1.255414058	-0.190584436	-1.96142254
High Social	48.617	7.588	7.807	7.031	-9.853204727	-10.2015597	-1.172788496	-0.363564946	-2.171024809
High Social	45.748	7.95	7.965	5.644	-11.92506599	-12.53651428	-1.096586282	-0.075809944	-2.180777237
Low Social	50.661	5.811	7.369	6.194	-12.20763524	-12.73648834	-1.114539475	-0.233160471	-2.184730759
High Social	51.861	5.6	8.131	8.564	-10.63426399	-11.72763824	-1.373022802	-0.164918556	-2.179943665
High Social	27.563	7.693	6.198	8.705	-13.21087646	-13.33796438	-1.276027163	-0.191251642	-2.063703145
High Social	53.679	6.396	12.322	11.838	-8.337731679	-9.373790105	-1.165037538	-0.260126134	-2.169312809

Table S2. SAS code for the data and analysis shown in Fig. 1.

```
DATA DataSetFig1; INPUT Group &$16. Corral &$16. CSF_OXT CSF_AVP Blot &$16. log10_p_PTEN_PTEN_
log10_p_AKT1_AKT1_; Lines;
High Social Q 46.761 6.22 Blot1 -0.0852885480858432 -2.17687152347855
Low Social Q 46.013 5.942 Blot1 -0.190604563237432 -2.22417635227326
High Social R 61.547 8.644 Blot2 -0.276778261581818 -2.2272118191264
High Social K 50.857 7.124 Blot2 -0.165349554207912 -2.17736189054273
Low Social R 86.68 3.152 Blot1 -0.299552308250423 -1.97098241186331
High Social L 41.029 5.206 Blot2 -0.295737826650851 -2.33062210150765
Low Social K 46.967 6.238 Blot1 -0.20932444620954 -2.01657088873934
Low Social K 43.645 6.691 Blot1 -0.131656353345104 -2.11312189410736
Low Social I 53.034 4.531 Blot2 -0.20784882228453 -2.05650007296392
Low Social M 39.24 6.135 Blot1 -0.234022352640254 -1.94841513248013
Low Social N 27.207 4.567 Blot2 -0.152687103440112 -2.15058971922419
Low Social M 84.82 6.678 Blot1 -0.0314382633339477 -1.97968337209127
High Social N 46.939 5.79 Blot1 -0.229728119926036 -1.94667450771242
High Social I 41.647 2.901 Blot1 -0.343990686864404 -2.14189502992807
High Social R 61.653 5.692 Blot1 -0.225806599417816 -2.03308649365278
High Social N 42.078 6.859 Blot2 -0.329361238707102 -2.1304181977994
Low Social N 37.623 6.036 Blot1 -0.125400565952727 -2.11891225535321
High Social N 59.596 7.43 Blot1 -0.181732505217158 -1.95850418778292
Low Social L 35.517 6.354 Blot2 -0.300877132259397 -2.11693002856273
Low Social H 40.439 8.229 Blot2 -0.21914103899142 -2.15833221606736
Low Social N 51.123 5.996 Blot2 -0.14694004626995 -2.2038235667791
High Social H 39.297 9.484 Blot1 -0.190584435525756 -1.9614225398245
High Social M 48.617 7.588 Blot2 -0.363564946172121 -2.17102480891455
High Social L 45.748 7.95 Blot2 -0.0758099439682486 -2.18077723738365
Low Social H 50.661 5.811 Blot2 -0.233160470564966 -2.18473075863214
High Social K 27.563 7.693 Blot1 -0.191251641768602 -2.06370314473609
High Social N 53.679 6.396 Blot2 -0.260126133563712 -2.16931280884789
;
RUN;

PROC GENMOD DATA=DataSetFig1 desc;
CLASS Corral Blot;
MODEL Group = Blot log10_p_PTEN_PTEN_ log10_p_AKT1_AKT1_ CSF_AVP CSF_OXT Corral/ DIST=Binomial LINK=Logit
ALPHA=0.05 type3;
RUN;
```

Table S3. SAS code for the data and analysis shown in Fig. 2.

```

DATA DataSetFig2; INPUT Group &$16. colony &$16. CSF_AVP csf_draw_time CaptureLatency Blot &$16.
log10_p_PTEN_PTEN_log10_p_AKT1_AKT1_; Lines;
High Social Q 6.22 31427 201 Blot1 -0.0852885480858432 -2.17687152347855
Low Social Q 5.942 31672 372 Blot1 -0.190604563237432 -2.22417635227326
High Social R 8.644 37259 303 Blot2 -0.276778261581818 -2.2272118191264
High Social K 7.124 31315 227 Blot2 -0.165349554207912 -2.17736189054273
Low Social R 3.152 33652 441 Blot1 -0.299552308250423 -1.97098241186331
High Social L 5.206 36419 692 Blot2 -0.295737826650851 -2.33062210150765
Low Social K 6.238 37201 834 Blot1 -0.20932444620954 -2.01657088873934
Low Social K 6.691 35124 306 Blot1 -0.131656353345104 -2.11312189410736
Low Social I 4.531 34980 159 Blot2 -0.20784882228453 -2.05650007296392
Low Social M 6.135 33068 217 Blot1 -0.234022352640254 -1.94841513248013
Low Social N 4.567 33275 257 Blot2 -0.152687103440112 -2.15058971922419
Low Social M 6.678 35319 419 Blot1 -0.0314382633339477 -1.97968337209127
High Social N 5.79 34694 177 Blot1 -0.229728119926036 -1.94667450771242
High Social I 2.901 31792 518 Blot1 -0.343990686864404 -2.14189502992807
High Social R 5.692 37163 416 Blot1 -0.225806599417816 -2.03308649365278
High Social N 6.859 36720 262 Blot2 -0.329361238707102 -2.1304181977994
Low Social N 6.036 34999 157 Blot1 -0.125400565952727 -2.11891225535321
Low Social N 7.487 36738 251 . . .
High Social N 7.43 34975 210 Blot1 -0.181732505217158 -1.95850418778292
Low Social L 6.354 37044 464 Blot2 -0.300877132259397 -2.11693002856273
Low Social H 8.229 31435 173 Blot2 -0.21914103899142 -2.15833221606736
Low Social N 5.996 37923 214 Blot2 -0.14694004626995 -2.2038235667791
Low Social M 5.084 33229 190 . . .
High Social H 9.484 31481 264 Blot1 -0.190584435525756 -1.9614225398245
High Social M 7.588 37032 443 Blot2 -0.363564946172121 -2.17102480891455
High Social L 7.95 34697 152 Blot2 -0.0758099439682486 -2.18077723738365
Low Social H 5.811 31403 195 Blot2 -0.233160470564966 -2.18473075863214
High Social P 5.6 33142 225 Blot2 -0.164918555635301 -2.17994366530132
High Social K 7.693 34969 212 Blot1 -0.191251641768602 -2.06370314473609
High Social N 6.396 33246 233 Blot2 -0.260126133563712 -2.16931280884789
;
RUN;

PROC GLM DATA=DataSetFig2 ALPHA=0.05 plots=none;
CLASS Group colony;
MODEL CSF_AVP = Group csf_draw_time colony CaptureLatency /ss3;

```

```
Lsmeans Group/stderr;  
RUN;  
PROC GLM DATA=DataSetFig2 ALPHA=0.05 plots=none;  
CLASS Group colony Blot;  
MODEL log10_p_PTEN_PTEN_ = Group colony Blot/ss3;  
Lsmeans Group/stderr;  
RUN;  
PROC GLM DATA=DataSetFig2 ALPHA=0.05 plots=none;  
CLASS Group colony Blot;  
MODEL log10_p_AKT1_AKT1_ = Group colony Blot/ss3;  
Lsmeans Group/stderr;  
RUN;
```

Table S4. SAS code for the data and analysis shown in Fig. 3A.

```
DATA DataSetFig3A; INPUT Corral &$16. CSF_AVP csf_draw_time CaptureLatency Sqrt_SumOfGR; Lines;  
Q 6.22 31427 201 8.77496438739212  
Q 5.942 31672 372 6.08276253029822  
R 8.644 37259 303 6  
K 7.124 31315 227 6.557438524302  
R 3.152 33652 441 3.74165738677394  
L 5.206 36419 692 5.3851648071345  
K 6.238 37201 834 6.78232998312527  
K 6.691 35124 306 4.69041575982343  
I 4.531 34980 159 1.73205080756888  
M 6.135 33068 217 5.74456264653803  
N 4.567 33275 257 2.82842712474619  
M 6.678 35319 419 5  
N 5.79 34694 177 6.40312423743285  
I 2.901 31792 518 3.16227766016838  
R 5.692 37163 416 4.24264068711929  
N 6.859 36720 262 7.61577310586391  
N 6.036 34999 157 5.09901951359278  
N 7.487 36738 251 2.82842712474619  
N 7.43 34975 210 6.85565460040104  
L 6.354 37044 464 2.64575131106459  
H 8.229 31435 173 4.47213595499958  
N 5.996 37923 214 1  
M 5.084 33229 190 2.64575131106459  
H 9.484 31481 264 4.89897948556636  
M 7.588 37032 443 5.47722557505166  
L 7.95 34697 152 6.70820393249937  
H 5.811 31403 195 3.60555127546399  
P 5.6 33142 225 4.35889894354067  
K 7.693 34969 212 3.46410161513775  
N 6.396 33246 233 4.69041575982343  
;  
RUN;  
  
PROC GLM DATA=DataSetFig3A;  
CLASS Corral;  
MODEL Sqrt_SumOfGR = csf_draw_time corral CaptureLatency CSF_AVP/ ss3;  
RUN;
```

Table S5. Confirmation of the specificity and validity of the statistical winnowing strategy for the monkey discovery cohort.

Biological Measure	Test Statistic	High-Social Monkey Group		Low-Social Monkey Group	
		LSM	95%CI	LSM	95%CI
CSF AVP level (pg/ml)	($F_{1,18} = 9.236$; $P = 0.0071$)	$6.86 \times 10^{+0}$	($6.28 \times 10^{+0}$, $7.44 \times 10^{+0}$)	$5.72 \times 10^{+0}$	($5.11 \times 10^{+0}$, $6.33 \times 10^{+0}$)
Blood AVP level (pg/ml)	($F_{1,18} = 1.736$; $P = 0.2042$)	$7.60 \times 10^{+0}$	($5.43 \times 10^{+0}$, $1.06 \times 10^{+1}$)	$1.01 \times 10^{+1}$	($7.10 \times 10^{+0}$, $1.45 \times 10^{+1}$)
Blood p-PTEN/PTEN ratio	($F_{1,17} = 1.089$; $P = 0.3113$)	5.98×10^{-1}	(5.33×10^{-1} , 6.72×10^{-1})	6.50×10^{-1}	(5.70×10^{-1} , 7.40×10^{-1})
Blood p-ERK/ERK ratio	($F_{1,17} = 0.980$; $P = 0.3361$)	5.93×10^{-2}	(4.49×10^{-2} , 7.83×10^{-2})	4.92×10^{-2}	(3.60×10^{-2} , 6.72×10^{-2})
CSF OXT level (pg/ml)	($F_{1,18} = 0.820$; $P = 0.3770$)	$4.77 \times 10^{+1}$	($3.84 \times 10^{+1}$, $5.70 \times 10^{+1}$)	$5.31 \times 10^{+1}$	($4.33 \times 10^{+1}$, $6.30 \times 10^{+1}$)
Blood AVPR _{V1a} mRNA level (- Δ CT)	($F_{1,18} = 0.487$; $P = 0.4941$)	$-1.15 \times 10^{+1}$	($-1.22 \times 10^{+1}$, $-1.08 \times 10^{+1}$)	$-1.18 \times 10^{+1}$	($-1.26 \times 10^{+1}$, $-1.11 \times 10^{+1}$)
Blood OXT level (pg/ml)	($F_{1,18} = 0.354$; $P = 0.5595$)	$7.43 \times 10^{+0}$	($4.86 \times 10^{+0}$, $9.99 \times 10^{+0}$)	$8.42 \times 10^{+0}$	($5.70 \times 10^{+0}$, $1.11 \times 10^{+1}$)
Blood p-AKT/AKT ratio	($F_{1,17} = 0.197$; $P = 0.6631$)	7.51×10^{-3}	(6.82×10^{-3} , 8.28×10^{-3})	7.74×10^{-3}	(6.94×10^{-3} , 8.63×10^{-3})
Blood OXTR mRNA level (- Δ CT)	($F_{1,18} = 0.018$; $P = 0.8953$)	$-1.07 \times 10^{+1}$	($-1.15 \times 10^{+1}$, $-9.96 \times 10^{+0}$)	$-1.08 \times 10^{+1}$	($-1.16 \times 10^{+1}$, $-9.98 \times 10^{+0}$)

Abbreviations: LSM, Least-Squares Means; CI, Confidence Interval; CSF, cerebrospinal fluid; AVP, arginine vasopressin; OXT, oxytocin. Differences between high-social and low-social monkey groups were tested using general linear models (GLM) as described in the main text. Measures are ordered by significance (note that only CSF AVP concentration is significant). The values are reported as LSM followed by 95%CI. Blood kinase ratios and blood AVP concentration were log-transformed for analysis to meet the requirements of GLM and are presented here as back-transformed values. Biological data were available from N=30 monkeys for the neuropeptide measures and N=28 monkeys for the kinase measures.

Table S6. SAS code for the data and analysis shown in Fig. 4.

```
DATA DataSetFig4; INPUT ObserverID &$16. group &$16. Mean_CSFDrawTime_ Mean_CaptureLatency_ Mean_CSF_AVP_
Lines;
OBS1 Low social 34497.5 112 5.117
OBS1 Low social 30923 128 8.745
OBS2 High social 31310 310 7.2515
OBS2 High social 31897.5 166.5 7.2235
OBS1 High social 36323 85 8.511
OBS2 Low social 32836 103 6.92
OBS3 High social 34816.5 134.5 10.4325
OBS1 Low social 31764 159 7.539
OBS2 High social 30949 147.5 7.056
OBS2 High social 16117 135.5 6.7575
OBS1 Low social 34758.5 195 6.456
OBS1 Low social 34686 151 7.1485
OBS2 High social 32035 221.5 7.2105
OBS3 Low social 33128.5 200 7.705
OBS1 Low social 30784.5 125 7.759
OBS3 Low social 32866 104 5.868
OBS3 Low social 30911.5 92 7.0875
OBS3 High social 37170.5 445 5.819
OBS3 Low social 30928 113 5.5625
OBS1 Low social 31942.5 167.5 7.6435
OBS1 Low social 33951.5 109 7.6525
OBS1 Low social 34584.5 158 8.7285
OBS3 High social 33076 176.5 8.909
OBS3 High social 36913.5 319 7.6125
OBS2 High social 31844 107.5 7.2065
OBS2 High social 33026 194 8.613
OBS1 Low social 34016 138 6.9255
OBS2 High social 31938.5 101 7.646
OBS2 High social 31962 128.5 7.8975
OBS3 High social 35023 260 9.25
;
RUN;

PROC GENMOD DATA=DataSetFig4 desc;
CLASS ObserverID;
```

```
MODEL group = ObserverID Mean_CSFDdrawTime_ Mean_CaptureLatency_ Mean_CSF_AVP_ / DIST=Binomial LINK=Logit
ALPHA=0.05 type3;
RUN;
PROC GLM DATA=DataSetFig4 ALPHA=0.05 plots=none;
CLASS ObserverID group;
MODEL Mean_CSF_AVP_ = ObserverID Mean_CSFDdrawTime_ Mean_CaptureLatency_ group/ss3;
LSMEANS Group / stderr;
RUN;
```

Table S7. SAS code for the data and analysis shown in Fig. 3B.

```
DATA DataSetICC; INPUT Monkey &$16. Timepoint CSF_AVP; Lines;
M01 1 23.178
M01 2 23.135
M01 3 34.264
M01 4 31.075
M02 1 15.433
M02 2 18.092
M02 3 27.243
M02 4 29.902
M03 1 33.925
M03 2 36.471
M03 3 49.045
M03 4 .
M04 1 14.463
M04 2 12.542
M04 3 16.947
M04 4 16.83
M05 1 25.423
M05 2 26.917
M05 3 43.73
M05 4 29.116
M06 1 17.153
M06 2 22.633
M06 3 21.844
M06 4 16.932
M07 1 22.964
M07 2 17.304
M07 3 24.332
M07 4 15.871
M08 1 16.046
M08 2 15.324
M08 3 28.218
M08 4 19.674
M09 1 21.128
M09 2 18.766
M09 3 23.55
M10 1 17.224
M10 2 17.994
```

```
M10 3 26.523  
M10 4 24.711
```

```
;  
RUN;
```

```
/*This PROC MIXED gives the variance components needed to figure the ICC*/  
/*see the SAS 14.3 online help for PROC MIXED, under the repeated measures example,*/  
/*Output 79.2.14 for details on figuring an ICC from this output*/
```

```
PROC MIXED DATA=DataSetICC;  
CLASS Monkey Timepoint;  
MODEL CSF_AVP = Timepoint ;  
repeated /type =cs subject= Monkey rcorr ;  
RUN;
```

```
/*This PROC MIXED forces SAS to calculate a p-value for the Monkey_ID,*/  
/*which gives a significance for the ICC*/
```

```
PROC MIXED DATA=DataSetICC;  
CLASS Monkey Timepoint;  
MODEL CSF_AVP = Timepoint Monkey ;  
LSMEANS Monkey;  
RUN;
```

Table S8. SAS code for the data and analysis shown in Fig. 5.

```
DATA DataSetFig5; INPUT Age_at_Sample_collection Group &$16. CSF_AVP SampleColTime; Lines;
6.58 ASD 32.969 0.395833333
6.08 ASD 11.466 0.416666667
6.75 ASD 24.086 0.326388889
18.83 ASD 30.115 0.486111111
9.92 ASD 19.765 0.469444444
5.3 ASD 12.071 0.402777778
6.92 ASD 15.802 0.525694444
6.33 CON 27.453 0.402777778
5.25 CON 26.415 0.353472222
7.83 CON 40.676 0.361111111
19.5 CON 41.408 0.768055556
10.42 CON 32.687 0.354166667
6.42 CON 26.805 0.361111111
7.5 CON 32.553 0.547916667
;
RUN;

PROC GENMOD DATA=DataSetFig5;
MODEL Group = CSF_AVP SampleColTime Age_at_Sample_collection/ DIST=Binomial LINK=Logit ALPHA=0.05 type3;
RUN;

PROC GLM DATA=DataSetFig5 ALPHA=0.05 plots=none;
CLASS group;
MODEL CSF_AVP = SampleColTime Age_at_Sample_collection group/ ss3;
lsmeans group/stderr;
RUN;
```