Supplement to Session 6

S6.2 Grouped data

S6.2.1 Observed and expected frequencies

```
* Set up formats for GCS at entry and outcome *;

proc format;
  value gcsfmt 1 = 'GCS 3-5'
               2 = 'GCS 6-8';
  value gosfmt 1 = 'Good'
               2 = 'Mod Dis'
               3 = 'Severe Dis'
               4 = 'Vegetative'
               5 = 'Dead';
run;

* Read in grouped data from severe head injury study (example 1) *;

data head1;
format gcsentry gcsfmt. gos gosfmt.;
input gcsentry gos frequency;
cards;
1  1   73
1  2   55
1  3   79
1  4   37
1  5  358
2  1  219
2  2  118
2  3   66
2  4   10
2  5   92
;
run;

* Fit proportional odds model for more favourable outcome with *
* GCS at entry and create dataset with original data, linear   *
* predictors and standard errors, and predicted cumulative     *
* probabilities                                                *;

proc logistic data=head1;
  weight frequency;
  class gcsentry (ref=’GCS 3-5’) / param=ref order=internal;
  model gos (order=internal) = gcsentry;
  output out=fittedcum xbeta=xbeta stdxbeta=stdxbeta
                       predprobs=cumulative;
run;
```

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept Good	$\hat{\alpha}_1$	1	**-2.1223**	0.1044	413.6292	<.0001
Intercept Mod Dis	$\hat{\alpha}_2$	1	**-1.2698**	0.0907	196.0272	<.0001
Intercept Severe Dis	$\hat{\alpha}_3$	1	**-0.6006**	0.0830	52.3620	<.0001
Intercept Vegetative	$\hat{\alpha}_4$	1	**-0.3861**	0.0816	22.3715	<.0001
gcsentry GCS 6-8	$\hat{\beta}$	1	**1.8984**	0.1210	246.1540	<.0001

GCS 3-5: $\hat{\eta}(0) = 0$ GCS 6-8: $\hat{\eta}(1) = 1.8984$

xbeta=xbeta stdxbeta=stdxbeta (Linear Predictors)

$\alpha_k + \hat{\eta}(0)$

Obs	gcsentry	gos	frequency	_LEVEL_	xbeta	stdxbeta
1	**GCS 3-5**	Good	73	Good	**-2.12226**	0.10435
2	**GCS 3-5**	Good	73	Mod Dis	**-1.26981**	0.09069
3	**GCS 3-5**	Good	73	Severe Dis	**-0.60063**	0.08300
4	**GCS 3-5**	Good	73	Vegetative	**-0.38610**	0.08163
5	GCS 3-5	Mod Dis	55	Good	-2.12226	0.10435
6	GCS 3-5	Mod Dis	55	Mod Dis	-1.26981	0.09069
7	GCS 3-5	Mod Dis	55	Severe Dis	-0.60063	0.08300
8	GCS 3-5	Mod Dis	55	Vegetative	-0.38610	0.08163
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
33	GCS 6-8	Vegetative	10	Good	-0.22381	0.08652
34	GCS 6-8	Vegetative	10	Mod Dis	0.62864	0.08882
35	GCS 6-8	Vegetative	10	Severe Dis	1.29782	0.09621
36	GCS 6-8	Vegetative	10	Vegetative	1.51235	0.09900
37	**GCS 6-8**	Dead	92	Good	**-0.22381**	0.08652
38	**GCS 6-8**	Dead	92	Mod Dis	**0.62864**	0.08882
39	**GCS 6-8**	Dead	92	Severe Dis	**1.29782**	0.09621
40	**GCS 6-8**	Dead	92	Vegetative	**1.51235**	0.09900

$\alpha_k + \hat{\eta}(1)$

`predprobs=cumulative` (Predicted Cumulative Probabilities)

$$\hat{Q}_1(0) = (1 + e^{-(-2.1222)})^{-1} \qquad \hat{Q}_2(0) = (1 + e^{-(-1.2698)})^{-1}$$

Obs	gcsentry	frequency	CP_Good	CP_Mod_ Dis	CP_Severe_ Dis	CP_Vegetative	CP_Dead
1	**GCS 3-5**	73	**0.10695**	**0.21929**	**0.35420**	**0.40466**	**1**
2	GCS 3-5	73	0.10695	0.21929	0.35420	0.40466	1
3	GCS 3-5	73	0.10695	0.21929	0.35420	0.40466	1
4	GCS 3-5	73	0.10695	0.21929	0.35420	0.40466	1
5	GCS 3-5	55	0.10695	0.21929	0.35420	0.40466	1
6	GCS 3-5	55	0.10695	0.21929	0.35420	0.40466	1
7	GCS 3-5	55	0.10695	0.21929	0.35420	0.40466	1
8	GCS 3-5	55	0.10695	0.21929	0.35420	0.40466	1
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
33	GCS 6-8	10	0.44428	0.65218	0.78547	0.81941	1
34	GCS 6-8	10	0.44428	0.65218	0.78547	0.81941	1
35	GCS 6-8	10	0.44428	0.65218	0.78547	0.81941	1
36	GCS 6-8	10	0.44428	0.65218	0.78547	0.81941	1
37	GCS 6-8	92	0.44428	0.65218	0.78547	0.81941	1
38	GCS 6-8	92	0.44428	0.65218	0.78547	0.81941	1
39	GCS 6-8	92	0.44428	0.65218	0.78547	0.81941	1
40	**GCS 6-8**	92	**0.44428**	**0.65218**	**0.78547**	**0.81941**	**1**

$$\hat{Q}_3(1) = (1 + e^{-(1.2978)})^{-1} \qquad \hat{Q}_4(1) = (1 + e^{-(1.5124)})^{-1} \qquad \hat{Q}_5(1)$$

```
* Fit proportional odds model for more favourable outcome with *
* GCS at entry and create dataset with original data, linear   *
* predictors and standard errors, and predicted probabilities  *;

proc logistic data=head1;
  weight frequency;
  class gcsentry (ref='GCS 3-5') / param=ref order=internal;
  model gos (order=internal) = gcsentry;
  output out=fittedind xbeta=xbeta stdxbeta=stdxbeta
                       predprobs=individual;
run;
```

predprobs=individual (Predicted Probabilities)

$$\hat{p}_1(0) = \hat{Q}_1(0) \qquad \hat{p}_2(0) = \hat{Q}_2(0) - \hat{Q}_1(0)$$

Obs	gcsentry	frequency	IP_Good	IP_Mod_ Dis	IP_Severe_ Dis	IP_Vegetative	IP_Dead
1	**GCS 3-5**	73	**0.10695**	**0.11234**	**0.13491**	**0.050457**	**0.59534**
2	GCS 3-5	73	0.10695	0.11234	0.13491	0.050457	0.59534
3	GCS 3-5	73	0.10695	0.11234	0.13491	0.050457	0.59534
4	GCS 3-5	73	0.10695	0.11234	0.13491	0.050457	0.59534
5	GCS 3-5	55	0.10695	0.11234	0.13491	0.050457	0.59534
6	GCS 3-5	55	0.10695	0.11234	0.13491	0.050457	0.59534
7	GCS 3-5	55	0.10695	0.11234	0.13491	0.050457	0.59534
8	GCS 3-5	55	0.10695	0.11234	0.13491	0.050457	0.59534
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
33	GCS 6-8	10	0.44428	0.20790	0.13329	0.033942	0.18059
34	GCS 6-8	10	0.44428	0.20790	0.13329	0.033942	0.18059
35	GCS 6-8	10	0.44428	0.20790	0.13329	0.033942	0.18059
36	GCS 6-8	10	0.44428	0.20790	0.13329	0.033942	0.18059
37	GCS 6-8	92	0.44428	0.20790	0.13329	0.033942	0.18059
38	GCS 6-8	92	0.44428	0.20790	0.13329	0.033942	0.18059
39	GCS 6-8	92	0.44428	0.20790	0.13329	0.033942	0.18059
40	**GCS 6-8**	92	**0.44428**	**0.20790**	**0.13329**	**0.033942**	**0.18059**

$$\hat{p}_3(1) = \hat{Q}_3(1) - \hat{Q}_2(1) \qquad \hat{p}_4(1) = \hat{Q}_4(1) - \hat{Q}_3(1) \qquad \hat{p}_5(1) = 1 - \hat{Q}_4(1)$$

Calculation of expected values

```
data obsexp;
set fitted1;

if _LEVEL_ = 1;

Obs_Good = frequency*(gos = 1);
Obs_Mod_Dis = frequency*(gos = 2);
Obs_Severe_Dis = frequency*(gos = 3);
Obs_Vegetative = frequency*(gos = 4);
Obs_Dead = frequency*(gos = 5);

Exp_Good = frequency*IP_Good;
Exp_Mod_Dis = frequency*IP_Mod_Dis;
Exp_Severe_Dis = frequency*IP_Severe_Dis;
Exp_Vegetative = frequency*IP_Vegetative;
Exp_Dead = frequency*IP_Dead;
run;

proc means data=obsexp sum;
  var Obs_Good Exp_Good Obs_Mod_Dis Exp_Mod_Dis Obs_Severe_Dis
  Exp_Severe_Dis Obs_Vegetative Exp_Vegetative Obs_Dead Exp_Dead;
  by gcsentry;
run;
```

```
------------------------------ gcsentry=GCS 3-5 ------------------------------

                              The MEANS Procedure

                         Variable                   Sum
                         ------------------------------
                         Obs_Good            73.0000000
                         Exp_Good            64.3851540
                         Obs_Mod_Dis         55.0000000
                         Exp_Mod_Dis         67.6271197
                         Obs_Severe_Dis      79.0000000
                         Exp_Severe_Dis      81.2162332
                         Obs_Vegetative      37.0000000
                         Exp_Vegetative      30.3753732
                         Obs_Dead           358.0000000
                         Exp_Dead           358.3961199
                         ------------------------------


------------------------------ gcsentry=GCS 6-8 ------------------------------

                         Variable                   Sum
                         ------------------------------
                         Obs_Good           219.0000000
                         Exp_Good           224.3613533
                         Obs_Mod_Dis        118.0000000
                         Exp_Mod_Dis        104.9897991
                         Obs_Severe_Dis      66.0000000
                         Exp_Severe_Dis      67.3102907
                         Obs_Vegetative      10.0000000
                         Exp_Vegetative      17.1405085
                         Obs_Dead            92.0000000
                         Exp_Dead            91.1980484
                         ------------------------------
```

S6.2.2 Score test for proportional odds

Score Test for the Proportional Odds Assumption

Chi-Square	DF	Pr > ChiSq
11.8845	**3**	**0.0078**

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	3125.079	2860.319
SC	3126.289	2861.832
-2 Log L	3117.079	**2850.319**

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	Good	1	**-2.1223**	0.1044	413.6292	<.0001
Intercept	Mod Dis	1	**-1.2698**	0.0907	196.0272	<.0001
Intercept	Severe Dis	1	**-0.6006**	0.0830	52.3620	<.0001
Intercept	Vegetative	1	**-0.3861**	0.0816	22.3715	<.0001
gcsentry	GCS 6-8	1	**1.8984**	0.1210	246.1540	<.0001

S6.3 Ungrouped data

Data from head injury study (example 2) in head2 dataset. See supplement to session 2 for details of these data.

```
/* age (yrs) continuous
   gcsmotor: 1 = None or extension
             2 = Abnormal or normal flexion
             3 = Localises
   treat: 0 = Control
          1 = Treated
   gos4: 1 = Good recovery
         2 = Moderate disability
         3 = Severe disability
         4 = Vegetative state or Dead
/*

* Fit proportional odds model for more favourable outcome with *
* age, GCS motor score and treatment and create dataset with   *
* original data, linear predictors and predicted probabilities *;

proc logistic data=head2;
  class gcsmotor (ref='1')
        treat (ref='0') / param=ref order=internal;
  model gos4 (order=internal) = age gcsmotor treat;
  output out=fitted2 xbeta=xbeta predprobs=individual;
run;

proc print data=fitted2 noobs;
  var patient age gcsmotor treat gos4 _LEVEL_ xbeta;
run;

* Calculate individual linear predictor for each patient using *
* estimate of alpha1 of -0.8217                                *;

data obsexp;
set fitted2;

if _LEVEL_ = 1;

eta = xbeta + 0.8217;

groupeta = (eta le -0.85) + (eta le -0.35) + (eta le 0.15) +
(eta le 0.65) + 1;

Obs1 = (gos4 = 1);
Obs2 = (gos4 = 2);
Obs3 = (gos4 = 3);
Obs4 = (gos4 = 4);

Exp1 = IP_1;
Exp2 = IP_2;
Exp3 = IP_3;
Exp4 = IP_4;
run;
```

```
proc print data=obsexp noobs;
  var patient age gcsmotor treat gos4 eta;
run;

proc sort data=obsexp;
  by groupeta;
run;

proc means data=obsexp sum nolabels;
  var Obs1 Exp1 Obs2 Exp2 Obs3 Exp3 Obs4 Exp4;
  by groupeta;
run;
```

Response Profile

Ordered Value	gos4	Total Frequency
1	1	113
2	2	57
3	3	60
4	4	111

Probabilities modeled are cumulated over the lower Ordered Values.

Class Level Information

Class	Value	Design Variables	
gcsmotor	1	0	0
	2	1	0
	3	0	1
treat	0	0	
	1	1	

Score Test for the Proportional Odds Assumption

Chi-Square	DF	Pr > ChiSq
5.4898	8	0.7042

Analysis of Maximum Likelihood Estimates

Parameter			DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	$\hat{\alpha}_1$	1	**-0.8217**	0.2924	7.8956	0.0050
Intercept	2	$\hat{\alpha}_2$	1	**0.00116**	0.2878	0.0000	0.9968
Intercept	3	$\hat{\alpha}_3$	1	**0.8461**	0.2914	8.4323	0.0037
age		$\hat{\beta}_1$	1	**-0.0247**	0.00682	13.0823	0.0003
gcsmotor	2	$\hat{\beta}_2$	1	**0.5058**	0.2324	4.7387	0.0295
gcsmotor	3	$\hat{\beta}_3$	1	**1.9026**	0.2832	45.1285	<.0001
treat	1	$\hat{\beta}_4$	1	**0.6019**	0.2054	8.5885	0.0034

S6.3.1 Observed and expected frequencies

From the model, each i[th] patient has:

- an individual linear predictor: $\eta(\underline{z}_i)$
- m fitted probabilities: $p_k(\underline{z}_i)$

SAS presents m – 1 rows for each patient in the output data set:

XBETA: $\alpha_k + \eta(\underline{z}_i)$ $k = 1,\ldots, m-1$

IP_1 – IP_4: $\hat{p}_1(\underline{z}_i)$ to $\hat{p}_4(\underline{z}_i)$

where

$$\hat{p}_k(\underline{z}_i) = \hat{Q}_k(\underline{z}_i) - \hat{Q}_{k-1}(\underline{z}_i)$$

$$\hat{Q}_k(\underline{z}_i) = (1+e^{-(\hat{\alpha}_k+\hat{\eta}(\underline{z}_i))})^{-1}$$

Linear Predictors

patient	age	gcsmotor	treat	gos4	_LEVEL_	xbeta
1	66	1	0	1	1	-2.44904
1	66	1	0	1	2	-1.62616
1	66	1	0	1	3	-0.78120
2	63	1	1	1	1	-1.77314
2	63	1	1	1	2	-0.95026
2	63	1	1	1	3	-0.10530
3	58	1	1	1	1	-1.64985
3	58	1	1	1	2	-0.82698
3	58	1	1	1	3	0.01798

etc.

from which $\hat{\eta}(\underline{z}_i)$ can be calculated:

Individual Linear Predictor

$$\hat{\eta}(\underline{z}_i) = X\hat{B}ETA - \hat{\alpha}_1$$

Observed category

patient	age	gcsmotor	treat	gos4	eta
1	66	1	0	1	-1.62734
2	63	1	1	1	-0.95144
3	58	1	1	1	-0.82815
4	54	1	0	1	-1.33146
5	51	1	1	1	-0.65556
6	49	1	0	1	-1.20818
7	49	1	1	1	-0.60625
8	46	1	0	1	-1.13421
9	46	1	1	1	-0.53228
10	43	1	0	1	-1.06024

etc.

```
--------------------------------- groupeta=1 ---------------------------------

                              The MEANS Procedure

                            Variable            Sum
                            ------------------------
                            Obs1         46.0000000
                            Exp1         46.8542492
                            Obs2         13.0000000
                            Exp2         11.7864356
                            Obs3          8.0000000
                            Exp3          7.7482055
                            Obs4          7.0000000
                            Exp4          7.6111096
                            ------------------------

--------------------------------- groupeta=2 ---------------------------------

                            Variable            Sum
                            ------------------------
                            Obs1         19.0000000
                            Exp1         16.8507484
                            Obs2          6.0000000
                            Exp2          8.8663366
                            Obs3          9.0000000
                            Exp3          7.9566853
                            Obs4         10.0000000
                            Exp4         10.3262297
                            ------------------------
```

```
--------------------------------- groupeta=3 ---------------------------------

                            Variable            Sum
                            ------------------------
                            Obs1         24.0000000
                            Exp1         25.3206838
                            Obs2         20.0000000
                            Exp2         16.5131673
                            Obs3         15.0000000
                            Exp3         17.2701098
                            Obs4         27.0000000
                            Exp4         26.8960391
                            ------------------------


--------------------------------- groupeta=4 ---------------------------------

                            Variable            Sum
                            ------------------------
                            Obs1         13.0000000
                            Exp1         16.8294428
                            Obs2         13.0000000
                            Exp2         13.5552428
                            Obs3         17.0000000
                            Exp3         17.1356130
                            Obs4         40.0000000
                            Exp4         35.4797013
                            ------------------------


--------------------------------- groupeta=5 ---------------------------------

                            Variable            Sum
                            ------------------------
                            Obs1         11.0000000
                            Exp1          6.7575371
                            Obs2          5.0000000
                            Exp2          6.4543096
                            Obs3         11.0000000
                            Exp3          9.8836514
                            Obs4         27.0000000
                            Exp4         30.9045018
                            ------------------------
```

Goodness-of-fit for the four category model
Observed and Expected Numbers of patients in each GOS Category in each Group

GroupGS	Good O	Good E	Moderate O	Moderate E	Severe O	Severe E	Veg/Dead O	Veg/Dead E	Total O	Total E
> 0.65	46	46.9	13	11.8	8	7.7	7	7.6	74	74.0
> 0.15 to <= 0.65	19	16.9	6	8.9	9	8.0	10	10.3	44	44.0
> -0.35 to <= 0.15	24	25.3	20	16.5	15	17.3	27	26.9	86	86.0
> -0.85 to <= -0.35	13	16.8	13	13.6	17	17.1	40	35.5	83	83.0
<= -0.85	11	6.8	5	6.5	11	9.9	27	30.9	54	54.0
Total	113	113.0	57	57.0	60	60.0	111	111.0	341	341.0

$$\chi^2 = \Sigma(O_{ks} - E_{ks})^2 / E_{ks} = 7.731$$

df = no. of cells – no. of parameters – no. of constraints

= (s groups × k categories)

– (q explanatory terms + no. of intercepts)

– (s groups)

= 20 – (4 + 3) – 5

= 8

c.f. χ^2_8, p = 0.46

(Ashby, Pocock and Shaper, 1986)

S6.4 To fit stratified proportional odds model using SAS PROC NLMIXED

```
* Read in grouped data from Alzheimer's disease example *;

data cgic;
input study treat cgic5 count;
cards;
1  1  1    4
1  1  2   23
1  1  3   45
1  1  4   22
1  1  5    2
1  0  1    2
1  0  2   22
1  0  3   54
1  0  4   29
1  0  5    3
2  1  1   14
2  1  2  119
2  1  3  180
2  1  4   54
2  1  5    6
2  0  1    1
2  0  2   22
2  0  3   35
2  0  4   11
2  0  5    3
3  1  1   13
3  1  2   20
3  1  3   24
3  1  4   10
3  1  5    1
3  0  1    7
3  0  2   16
3  0  3   17
3  0  4   10
3  0  5    3
4  1  1   21
4  1  2  106
4  1  3  175
4  1  4   62
4  1  5   17
4  0  1    8
4  0  2   24
4  0  3   73
4  0  4   52
4  0  5   13
5  1  1    3
5  1  2   14
5  1  3   19
5  1  4    3
5  1  5    0
5  0  1    2
5  0  2   13
```

```
5  0  3   18
5  0  4    7
5  0  5    1
;
run;

* Macro for creating intercept terms in stratified model *;

%macro alpha(i,k);
  %if &k=1 %then a&i.1;
  %if &k=2 %then a&i.1 + a&i.2;
  %if &k=3 %then a&i.1 + a&i.2 + a&i.3;
  %if &k=4 %then a&i.1 + a&i.2 + a&i.3 + a&i.4;
%mend alpha;

* Fit proportional odds model for more favourable outcome with *
* treatment, stratified by study                               *;

* Use PROC NLMIXED with study and treatment as fixed effects *;

proc nlmixed data=cgic;
  parms  a11 a12 a13 a14 a21 a22 a23 a24 a31 a32 a33 a34
         a41 a42 a43 a44 a51 a52 a53 a54 = 1, beta1 = 0;
  bounds a12 a13 a14 a22 a23 a24 a32 a33 a34 a42 a43 a44
         a52 a53 a54 > 0;

  eta = beta1*treat;

  if study = 1 and cgic5 = 1 then do;
    qk = 1/(1+exp(-(%alpha(1,1))-eta));
    qk_1 = 0;
  end;
  if study = 1 and cgic5 = 2 then do;
    qk = 1/(1+exp(-(%alpha(1,2))-eta));
    qk_1 = 1/(1+exp(-(%alpha(1,1))-eta));
  end;
  if study = 1 and cgic5 = 3 then do;
    qk = 1/(1+exp(-(%alpha(1,3))-eta));
    qk_1 = 1/(1+exp(-(%alpha(1,2))-eta));
  end;
  if study = 1 and cgic5 = 4 then do;
    qk = 1/(1+exp(-(%alpha(1,4))-eta));
    qk_1 = 1/(1+exp(-(%alpha(1,3))-eta));
  end;
  if study = 1 and cgic5 = 5 then do;
    qk = 1;
    qk_1 = 1/(1+exp(-(%alpha(1,4))-eta));
  end;
  if study = 2 and cgic5 = 1 then do;
    qk = 1/(1+exp(-(%alpha(2,1))-eta));
    qk_1 = 0;
  end;
  if study = 2 and cgic5 = 2 then do;
    qk = 1/(1+exp(-(%alpha(2,2))-eta));
    qk_1 = 1/(1+exp(-(%alpha(2,1))-eta));
  end;
```

```
if study = 2 and cgic5 = 3 then do;
  qk = 1/(1+exp(-(%alpha(2,3))-eta));
  qk_1 = 1/(1+exp(-(%alpha(2,2))-eta));
end;
if study = 2 and cgic5 = 4 then do;
  qk = 1/(1+exp(-(%alpha(2,4))-eta));
  qk_1 = 1/(1+exp(-(%alpha(2,3))-eta));
end;
if study = 2 and cgic5 = 5 then do;
  qk = 1;
  qk_1 = 1/(1+exp(-(%alpha(2,4))-eta));
end;
if study = 3 and cgic5 = 1 then do;
  qk = 1/(1+exp(-(%alpha(3,1))-eta));
  qk_1 = 0;
end;
if study = 3 and cgic5 = 2 then do;
  qk = 1/(1+exp(-(%alpha(3,2))-eta));
  qk_1 = 1/(1+exp(-(%alpha(3,1))-eta));
end;
if study = 3 and cgic5 = 3 then do;
  qk = 1/(1+exp(-(%alpha(3,3))-eta));
  qk_1 = 1/(1+exp(-(%alpha(3,2))-eta));
end;
if study = 3 and cgic5 = 4 then do;
  qk = 1/(1+exp(-(%alpha(3,4))-eta));
  qk_1 = 1/(1+exp(-(%alpha(3,3))-eta));
end;
if study = 3 and cgic5 = 5 then do;
  qk = 1;
  qk_1 = 1/(1+exp(-(%alpha(3,4))-eta));
end;
if study = 4 and cgic5 = 1 then do;
  qk = 1/(1+exp(-(%alpha(4,1))-eta));
  qk_1 = 0;
end;
if study = 4 and cgic5 = 2 then do;
  qk = 1/(1+exp(-(%alpha(4,2))-eta));
  qk_1 = 1/(1+exp(-(%alpha(4,1))-eta));
end;
if study = 4 and cgic5 = 3 then do;
  qk = 1/(1+exp(-(%alpha(4,3))-eta));
  qk_1 = 1/(1+exp(-(%alpha(4,2))-eta));
end;
if study = 4 and cgic5 = 4 then do;
  qk = 1/(1+exp(-(%alpha(4,4))-eta));
  qk_1 = 1/(1+exp(-(%alpha(4,3))-eta));
end;
if study = 4 and cgic5 = 5 then do;
  qk = 1;
  qk_1 = 1/(1+exp(-(%alpha(4,4))-eta));
end;
if study = 5 and cgic5 = 1 then do;
  qk = 1/(1+exp(-(%alpha(5,1))-eta));
  qk_1 = 0;
end;
```

```
  if study = 5 and cgic5 = 2 then do;
    qk = 1/(1+exp(-(%alpha(5,2))-eta));
    qk_1 = 1/(1+exp(-(%alpha(5,1))-eta));
  end;
  if study = 5 and cgic5 = 3 then do;
    qk = 1/(1+exp(-(%alpha(5,3))-eta));
    qk_1 = 1/(1+exp(-(%alpha(5,2))-eta));
  end;
  if study = 5 and cgic5 = 4 then do;
    qk = 1/(1+exp(-(%alpha(5,4))-eta));
    qk_1 = 1/(1+exp(-(%alpha(5,3))-eta));
  end;
  if study = 5 and cgic5 = 5 then do;
    qk = 1;
    qk_1 = 1/(1+exp(-(%alpha(5,4))-eta));
  end;

  p=qk-qk_1;

  if p>1e-8 then ll=log(p);
  else ll=-1e100;

  model cgic5 ~ general(ll);
  replicate count;
run;
```

Proportional odds model for more favourable outcome with treatment, stratified by study

```
                         The NLMIXED Procedure

                            Specifications

      Data Set                                      WORK.CGIC
      Dependent Variable                            cgic5
      Distribution for Dependent Variable           General
      Replicate Variable                            count
      Optimization Technique                        Dual Quasi-Newton
      Integration Method                            None


                              Dimensions

                  Observations Used                        49
                  Observations Not Used                     1
                  Total Observations                       50
                  Parameters                               21


                              Parameters

    a11       a12       a13       a14       a21       a22       a23       a24

      1         1         1         1         1         1         1         1

                              Parameters

    a31       a32       a33       a34       a41       a42       a43       a44

      1         1         1         1         1         1         1         1
```

Parameters

a51	a52	a53	a54	beta1	NegLogLike
1	1	1	1	0	3542.5429

Iteration History

Iter	Calls	NegLogLike	Diff	MaxGrad	Slope
1	3	2524.04195	1018.501	282.1493	-9029.73
2	4	2339.70038	184.3416	77.06235	-290.422
3	6	2148.73504	190.9653	81.13803	-158.602
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
46	76	1779.83463	0.000045	0.023433	-0.00023
47	77	1779.83456	0.00007	0.019578	-0.00015
48	79	1779.83456	7.123E-6	0.023019	-0.00001

NOTE: GCONV convergence criterion satisfied.

Fit Statistics

-2 Log Likelihood	**3559.7**
AIC (smaller is better)	3601.7
AICC (smaller is better)	3635.9
BIC (smaller is better)	3641.4

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower
a11	-3.7588	0.4182	49	-8.99	<.0001	0.05	-4.5991
a12	2.4037	0.3966	49	6.06	<.0001	0.05	1.6067
a13	2.1187	0.1830	49	11.58	<.0001	0.05	1.7509
a14	2.7218	0.4407	49	6.18	<.0001	0.05	1.8361
a21	-3.7880	0.2802	49	-13.52	<.0001	0.05	-4.3511
a22	2.7446	0.2563	49	10.71	<.0001	0.05	2.2296
a23	2.2388	0.1339	49	16.72	<.0001	0.05	1.9697
a24	2.2737	0.3199	49	7.11	<.0001	0.05	1.6309
a31	-1.9186	0.2548	49	-7.53	<.0001	0.05	-2.4307
a32	1.4804	0.2257	49	6.56	<.0001	0.05	1.0269
a33	1.5588	0.2180	49	7.15	<.0001	0.05	1.1206
a34	1.9931	0.4752	49	4.19	0.0001	0.05	1.0381
a41	-3.2719	0.2096	49	-15.61	<.0001	0.05	-3.6930
a42	2.0010	0.1793	49	11.16	<.0001	0.05	1.6406
a43	1.9762	0.1083	49	18.24	<.0001	0.05	1.7585
a44	1.8347	0.1741	49	10.54	<.0001	0.05	1.4849
a51	-2.9833	0.4672	49	-6.39	<.0001	0.05	-3.9222
a52	2.3166	0.4475	49	5.18	<.0001	0.05	1.4173
a53	2.2695	0.3333	49	6.81	<.0001	0.05	1.5996
a54	2.5480	0.9674	49	2.63	0.0113	0.05	0.6040
beta1 $\hat{\theta}$	**0.5048**	**0.1121**	49	4.50	<.0001	0.05	0.2795

$\hat{\theta}$ is estimate of the log odds of a more favourable outcome with Tacrine compared to placebo, stratified by study.

S6.5 Analysis of grouped ordinal data using R

```
# Read grouped data from severe head injury study (example 1) into a data frame

head1 <- data.frame(GCSENTRY=c(1,1,1,1,1,2,2,2,2,2),
                    GOS=c(1,2,3,4,5,1,2,3,4,5),
                    FREQUENCY=c(73,55,79,37,358,219,118,66,10,92))


# Convert ordinal outcome variable from numeric variable to factor variable
# with 5 (Dead) as the first level and 1 (Good) as the last level so that
# lrm function will model cumulative probabilities of more favourable outcome

head1$GOS <- factor(head1$GOS, levels=c('5','4','3','2','1'),
labels=c('Dead','Vegetative','Severe Dis','Mod Dis','Good'))


# Convert categorical predictor from numeric variable to factor variable

head1$GCSENTRY <- factor(head1$GCSENTRY, levels=c('1','2'),
labels=c('GCS 3-5','GCS 6-8'))


# Load rms package

library(rms)


# Fit proportional odds model for more favourable outcome with GCS at entry

pom.fit1 <- lrm(GOS ~ GCSENTRY, data=head1, weight=FREQUENCY)
print(pom.fit1)
```

The **weight=** option is used because the data are in summarised form rather than with a row for each individual.

```
# Obtain linear predictors, predicted cumulative probabilities, and
# predicted probabilities

# Linear predictors for log odds of GOS = 'Good', GOS = 'Mod Dis' or better,
# GOS = 'Severe Dis' or better, and GOS = 'Vegetative' or better are derived
# using fourth, third, second, and first intercepts respectively

fitted1 <- cbind(head1,
                 XBETA1=predict(pom.fit1, head1, type='lp', kint=4),
                 XBETA2=predict(pom.fit1, head1, type='lp', kint=3),
                 XBETA3=predict(pom.fit1, head1, type='lp', kint=2),
                 XBETA4=predict(pom.fit1, head1, type='lp', kint=1),
                 predict(pom.fit1, head1, type='fitted'),
                 predict(pom.fit1, head1, type='fitted.ind'))

print(fitted1[,c('GCSENTRY','GOS','FREQUENCY','XBETA1','XBETA2','XBETA3',
'XBETA4')])
print(fitted1[,c('GCSENTRY','y>=Good','y>=Mod Dis','y>=Severe Dis',
'y>=Vegetative')])
print(fitted1[,c('GCSENTRY','GOS=Good','GOS=Mod Dis','GOS=Severe Dis',
'GOS=Vegetative','GOS=Dead')])
```

```
# Calculate expected values for each GOS category by GCS at entry

expected <- data.frame(GCSENTRY=head1$GCSENTRY,
                       E_Good=fitted1$FREQUENCY*fitted1$'GOS=Good',
                       E_Mod_Dis=fitted1$FREQUENCY*fitted1$'GOS=Mod Dis',
                       E_Severe_Dis=fitted1$FREQUENCY*fitted1$'GOS=Severe Dis',
                       E_Vegetative=fitted1$FREQUENCY*fitted1$'GOS=Vegetative',
                       E_Dead=fitted1$FREQUENCY*fitted1$'GOS=Dead')

EXPECTED <- c(sum(expected[expected$GCSENTRY=='GCS 3-5','E_Good']),
              sum(expected[expected$GCSENTRY=='GCS 3-5','E_Mod_Dis']),
              sum(expected[expected$GCSENTRY=='GCS 3-5','E_Severe_Dis']),
              sum(expected[expected$GCSENTRY=='GCS 3-5','E_Vegetative']),
              sum(expected[expected$GCSENTRY=='GCS 3-5','E_Dead']),
              sum(expected[expected$GCSENTRY=='GCS 6-8','E_Good']),
              sum(expected[expected$GCSENTRY=='GCS 6-8','E_Mod_Dis']),
              sum(expected[expected$GCSENTRY=='GCS 6-8','E_Severe_Dis']),
              sum(expected[expected$GCSENTRY=='GCS 6-8','E_Vegetative']),
              sum(expected[expected$GCSENTRY=='GCS 6-8','E_Dead']))

obsexp <- data.frame(head1, EXPECTED)

print(obsexp)
```

Proportional Odds Model

```
Logistic Regression Model

lrm(formula = GOS ~ GCSENTRY, data = head1, weights = FREQUENCY)

Frequencies of Responses

      Dead Vegetative Severe Dis    Mod Dis       Good
         2          2          2          2          2

Sum of Weights by Response Category

  0   1   2   3   4
450  47 145 173 292

                       Model Likelihood     Discrimination    Rank Discrim.
                          Ratio Test           Indexes           Indexes

Obs              10    LR chi2     266.76    R2        0.228    C       0.681
Sum of weights 1107    d.f.             1    g         1.055    Dxy     0.361
max |deriv|   1e-07    Pr(> chi2) <0.0001    gr        2.871    gamma   0.658
                                             gp        0.240    tau-a   0.261
                                             Brier     0.196
```

		Coef	S.E.	Wald Z	Pr(>|Z|)
y>=Vegetative	$\hat{\alpha}_4$	**-0.3861**	0.0816	-4.73	<0.0001
y>=Severe Dis	$\hat{\alpha}_3$	**-0.6006**	0.0831	-7.23	<0.0001
y>=Mod Dis	$\hat{\alpha}_2$	**-1.2698**	0.0915	-13.88	<0.0001
y>=Good	$\hat{\alpha}_1$	**-2.1223**	0.1051	-20.19	<0.0001
GCSENTRY=GCS 6-8	$\hat{\beta}$	**1.8985**	0.1214	15.64	<0.0001

Linear Predictors

	GCSENTRY	GOS	FREQUENCY	XBETA1	XBETA2	XBETA3	XBETA4
1	**GCS 3-5**	Good	73	**-2.1222861**	**-1.2698372**	**-0.6006347**	**-0.3861009**
2	GCS 3-5	Mod Dis	55	-2.1222861	-1.2698372	-0.6006347	-0.3861009
3	GCS 3-5	Severe Dis	79	-2.1222861	-1.2698372	-0.6006347	-0.3861009
4	GCS 3-5	Vegetative	37	-2.1222861	-1.2698372	-0.6006347	-0.3861009
5	GCS 3-5	Dead	358	-2.1222861	-1.2698372	-0.6006347	-0.3861009
6	GCS 6-8	Good	219	-0.2238070	0.6286419	1.2978444	1.5123782
7	GCS 6-8	Mod Dis	118	-0.2238070	0.6286419	1.2978444	1.5123782
8	GCS 6-8	Severe Dis	66	-0.2238070	0.6286419	1.2978444	1.5123782
9	GCS 6-8	Vegetative	10	-0.2238070	0.6286419	1.2978444	1.5123782
10	**GCS 6-8**	Dead	92	**-0.2238070**	**0.6286419**	**1.2978444**	**1.5123782**

Predicted Cumulative Probabilities

	GCSENTRY	y>=Good	y>=Mod Dis	y>=Severe Dis	y>=Vegetative
1	**GCS 3-5**	**0.1069495**	**0.2192851**	**0.3541985**	**0.4046563**
2	GCS 3-5	0.1069495	0.2192851	0.3541985	0.4046563
3	GCS 3-5	0.1069495	0.2192851	0.3541985	0.4046563
4	GCS 3-5	0.1069495	0.2192851	0.3541985	0.4046563
5	GCS 3-5	0.1069495	0.2192851	0.3541985	0.4046563
6	GCS 6-8	0.4442806	0.6521815	0.7854720	0.8194134
7	GCS 6-8	0.4442806	0.6521815	0.7854720	0.8194134
8	GCS 6-8	0.4442806	0.6521815	0.7854720	0.8194134
9	GCS 6-8	0.4442806	0.6521815	0.7854720	0.8194134
10	**GCS 6-8**	**0.4442806**	**0.6521815**	**0.7854720**	**0.8194134**

Predicted Probabilities

	GCSENTRY	GOS=Good	GOS=Mod Dis	GOS=Severe Dis	GOS=Vegetative	GOS=Dead
1	**GCS 3-5**	**0.1069495**	**0.1123356**	**0.1349134**	**0.05045776**	**0.5953437**
2	GCS 3-5	0.1069495	0.1123356	0.1349134	0.05045776	0.5953437
3	GCS 3-5	0.1069495	0.1123356	0.1349134	0.05045776	0.5953437
4	GCS 3-5	0.1069495	0.1123356	0.1349134	0.05045776	0.5953437
5	GCS 3-5	0.1069495	0.1123356	0.1349134	0.05045776	0.5953437
6	GCS 6-8	0.4442806	0.2079008	0.1332905	0.03394140	0.1805866
7	GCS 6-8	0.4442806	0.2079008	0.1332905	0.03394140	0.1805866
8	GCS 6-8	0.4442806	0.2079008	0.1332905	0.03394140	0.1805866
9	GCS 6-8	0.4442806	0.2079008	0.1332905	0.03394140	0.1805866
10	**GCS 6-8**	**0.4442806**	**0.2079008**	**0.1332905**	**0.03394140**	**0.1805866**

Observed and Expected Frequencies

	GCSENTRY	GOS	FREQUENCY	EXPECTED
1	GCS 3-5	Good	73	64.38362
2	GCS 3-5	Mod Dis	55	67.62603
3	GCS 3-5	Severe Dis	79	81.21786
4	GCS 3-5	Vegetative	37	30.37557
5	GCS 3-5	Dead	358	358.39692
6	GCS 6-8	Good	219	224.36173
7	GCS 6-8	Mod Dis	118	104.98991
8	GCS 6-8	Severe Dis	66	67.31172
9	GCS 6-8	Vegetative	10	17.14041
10	GCS 6-8	Dead	92	91.19624

S6.6 Analysis of ungrouped ordinal data using R

Data from head injury study (example 2) in head2 data frame. See supplement to session 2 for details of these data.

```
# Load saved workspace which contains data frame with ungrouped data from
# head injury study (example 2)

load('head2.RData')


# Convert ordinal outcome variable from numeric variable to factor variable
# with 4 (Veg/Dead) as the first level and 1 (Good) as the last level so that
# lrm function will model cumulative probabilities of more favourable outcome

head2$GOS4 <- factor(head2$GOS4, levels=c('4','3','2','1'))


# Convert categorical predictors from numeric variables to factor variables

head2$GCSMOTOR <- factor(head2$GCSMOTOR, levels=c('1','2','3'))
head2$TREAT <- factor(head2$TREAT, levels=c('0','1'))


# Load rms package

library(rms)


# Fit proportional odds model for more favourable outcome with age,
# GCS motor score and treatment

pom.fit2 <- lrm(GOS4 ~ AGE + GCSMOTOR + TREAT, data=head2)
print(pom.fit2)


# Obtain linear predictors, predicted cumulative probabilities, and
# predicted probabilities

# Linear predictors for log odds of GOS4 = 1 or better, GOS4 = 2 or better,
# and GOS4 = 3 or better are derived using third, second, and first intercepts
# respectively

fitted2 <- cbind(head2,
                 XBETA1=predict(pom.fit2, head2, type='lp', kint=3),
                 XBETA2=predict(pom.fit2, head2, type='lp', kint=2),
                 XBETA3=predict(pom.fit2, head2, type='lp', kint=1),
                 predict(pom.fit2, head2, type='fitted'),
                 predict(pom.fit2, head2, type='fitted.ind'))

print(fitted2[,c('PATIENT','AGE','GCSMOTOR','TREAT','GOS4','XBETA1','XBETA2',
'XBETA3')], row.names=FALSE)


# Calculate individual linear predictor for each patient

fitted2$ETA <- fitted2$XBETA1 - pom.fit2$coefficients['y>=1']

print(fitted2[,c('PATIENT',’AGE’,'GCSMOTOR','TREAT','GOS4','ETA')],
row.names=FALSE)
```

```
# Classify each patient into one of five groups based on value of individual
# linear predictor

fitted2$GROUPETA <- 6 - cut(fitted2$ETA, breaks=c(-Inf,-0.85,-0.35,0.15,
0.65,Inf), labels=FALSE)


# Calculate observed frequencies for each GOS4 category by individual linear
# predictor group

OBS1 <- rowsum(as.numeric(fitted2$GOS4==1), fitted2$GROUPETA)
OBS2 <- rowsum(as.numeric(fitted2$GOS4==2), fitted2$GROUPETA)
OBS3 <- rowsum(as.numeric(fitted2$GOS4==3), fitted2$GROUPETA)
OBS4 <- rowsum(as.numeric(fitted2$GOS4==4), fitted2$GROUPETA)


# Calculate expected frequencies for each GOS4 category by individual linear
# predictor group

EXP1 <- rowsum(fitted2$'GOS4=1', fitted2$GROUPETA)
EXP2 <- rowsum(fitted2$'GOS4=2', fitted2$GROUPETA)
EXP3 <- rowsum(fitted2$'GOS4=3', fitted2$GROUPETA)
EXP4 <- rowsum(fitted2$'GOS4=4', fitted2$GROUPETA)

GROUPETA <- c(1:5)
GROUPETA <- factor(GROUPETA, levels=c('1','2','3','4','5'),
labels=c('> 0.65','> 0.15 to <= 0.65','>-0.35 to <= 0.15',
'>-0.85 to <=-0.35','<=-0.85'))

obsexp <- data.frame(GROUPETA,OBS1,EXP1,OBS2,EXP2,OBS3,EXP3,OBS4,EXP4)
print(obsexp, row.names=FALSE)
```

Proportional Odds Model

```
Logistic Regression Model

lrm(formula = GOS4 ~ AGE + GCSMOTOR + TREAT, data = head2)

Frequencies of Responses

  4   3   2   1
111  60  57 113

                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test            Indexes          Indexes

Obs            341    LR chi2      68.02    R2       0.194    C       0.682
max |deriv| 7e-10    d.f.             4    g        0.987    Dxy     0.365
                      Pr(> chi2) <0.0001    gr       2.683    gamma   0.367
                                            gp       0.216    tau-a   0.265
                                            Brier    0.213
```

		Coef	S.E.	Wald Z	Pr(>|Z|)
y>=3	$\hat{\alpha}_3$	**0.8462**	0.2918	2.90	0.0037
y>=2	$\hat{\alpha}_2$	**0.0012**	0.2873	0.00	0.9967
y>=1	$\hat{\alpha}_1$	**-0.8217**	0.2919	-2.82	0.0049
AGE	$\hat{\beta}_1$	**-0.0247**	0.0069	-3.56	0.0004
GCSMOTOR=2	$\hat{\beta}_2$	**0.5058**	0.2338	2.16	0.0305
GCSMOTOR=3	$\hat{\beta}_3$	**1.9026**	0.2803	6.79	<0.0001
TREAT=1	$\hat{\beta}_4$	**0.6019**	0.2056	2.93	0.0034

Linear Predictors

PATIENT	AGE	GCSMOTOR	TREAT	GOS4	XBETA1	XBETA2	XBETA3
1	66	1	0	1	-2.449087433	-1.6262100113	-0.781243041
2	63	1	1	1	-1.773176425	-0.9502990033	-0.105332033
3	58	1	1	1	-1.649888539	-0.8270111168	0.017955854
4	54	1	0	1	-2.153196506	-1.3303190836	-0.485352113
5	51	1	1	1	-1.477285498	-0.6544080756	0.190558895
6	49	1	0	1	-2.029908619	-1.2070311971	-0.362064227
7	49	1	1	1	-1.427970343	-0.6050929210	0.239874049
8	46	1	0	1	-1.955935887	-1.1330584651	-0.288091495
9	46	1	1	1	-1.353997611	-0.5311201891	0.313846781
10	43	1	0	1	-1.881963155	-1.0590857332	-0.214118763

etc.

Individual Linear Predictor

PATIENT	AGE	GCSMOTOR	TREAT	GOS4	ETA
1	66	1	0	1	-1.627400102
2	63	1	1	1	-0.951489094
3	58	1	1	1	-0.828201208
4	54	1	0	1	-1.331509175
5	51	1	1	1	-0.655598167
6	49	1	0	1	-1.208221288
7	49	1	1	1	-0.606283012
8	46	1	0	1	-1.134248556
9	46	1	1	1	-0.532310280
10	43	1	0	1	-1.060275824

etc.

Observed and Expected Frequencies

GROUPETA	OBS1	EXP1	OBS2	EXP2	OBS3	EXP3	OBS4	EXP4
> 0.65	46	46.854106	13	11.786505	8	7.748265	7	7.611123
> 0.15 to <= 0.65	19	16.850622	6	8.866377	9	7.956749	10	10.326252
>-0.35 to <= 0.15	24	25.320507	20	16.513212	15	17.270218	27	26.896063
>-0.85 to <=-0.35	13	16.829245	13	13.555228	17	17.135697	40	35.479830
<=-0.85	11	6.757345	5	6.454214	11	9.883618	27	30.904824

S6.7 Stratified proportional odds model using R

```
# Read grouped data from Alzheimer's disease example into a data frame

cgic <- expand.grid(CGIC5=c(1:5), TREAT=c(1,0), STUDY=c(1:5))
cgic$COUNT <- c(4,23,45,22,2,
                2,22,54,29,3,
                14,119,180,54,6,
                1,22,35,11,3,
                13,20,24,10,1,
                7,16,17,10,3,
                21,106,175,62,17,
                8,24,73,52,13,
                3,14,19,3,0,
                2,13,18,7,1)


# Convert ordinal outcome variable from numeric variable to factor variable with
# 1 (Very much or much improved) as first and 5 (Very much or much worse) as
# last level so that clm function will model cumulative probabilities of more
# favourable outcome

cgic$CGIC5 <- factor(cgic$CGIC5, levels=c('1','2','3','4','5'))


# Convert stratification variable from numeric variable to factor variable

cgic$STUDY <- factor(cgic$STUDY, levels=c('1','2','3','4','5'))


# Convert categorical predictor from numeric variable to factor variable

cgic$TREAT <- factor(cgic$TREAT, levels=c('0','1'),
labels=c('Placebo','Tacrine'))


# load ordinal package

library (ordinal)


# Fit proportional odds model for more favourable outcome with treatment,
# stratified by study

clm.fit1 <- clm(CGIC5 ~ TREAT, nominal = ~ STUDY, data=cgic, weights=COUNT,
link='logistic')
summary(clm.fit1)


# Fit model for more favourable outcome stratified by study but without
# treatment

clm.fit2 <- clm(CGIC5 ~ 1, nominal = ~ STUDY, data=cgic, weights=COUNT,
link='logistic')
summary(clm.fit2)


# Likelihood ratio test for treatment, stratified by study, from taking
# difference between -2 Log L for last two clm function calls

anova(clm.fit2, clm.fit1)
```

Proportional odds model for more favourable outcome with treatment, stratified by study

```
Call:
clm(location = CGIC5 ~ TREAT, nominal = ~STUDY, data = cgic,
    weights = COUNT, link = "logistic")

Location coefficients:
                   Estimate Std. Error z value Pr(>|z|)
TREATTacrine   β̂   -0.5048   0.1121    -4.5031 6.6964e-06

No scale coefficients

Threshold coefficients:
                Estimate Std. Error z value
1|2.(Intercept) -3.7624   0.4189    -8.9821
2|3.(Intercept) -1.3553   0.1716    -7.8996
3|4.(Intercept)  0.7637   0.1644     4.6450
4|5.(Intercept)  3.4810   0.4552     7.6481
1|2.STUDY2      -0.0258   0.4923    -0.0524
2|3.STUDY2       0.3119   0.1944     1.6040
3|4.STUDY2       0.4318   0.2073     2.0832
4|5.STUDY2      -0.0131   0.5663    -0.0231
1|2.STUDY3       1.8436   0.4819     3.8255
2|3.STUDY3       0.9171   0.2450     3.7435
3|4.STUDY3       0.3567   0.2780     1.2832
4|5.STUDY3      -0.3684   0.6814    -0.5406
1|2.STUDY4       0.4904   0.4572     1.0726
2|3.STUDY4       0.0844   0.1896     0.4451
3|4.STUDY4      -0.0584   0.1868    -0.3125
4|5.STUDY4      -0.9410   0.4910    -1.9166
1|2.STUDY5       0.7808   0.6211     1.2571
2|3.STUDY5       0.6889   0.2812     2.4498
3|4.STUDY5       0.8384   0.3616     2.3183
4|5.STUDY5       0.6702   1.1038     0.6072

log-likelihood: -1779.834
AIC: 3601.669
Condition number of Hessian: 969.5578
```

Note: The clm function interprets a response variable with more than two categories as an ordinal response and models the cumulative probabilities of lower ordered values. So in this example to model the cumulative probabilities of more favourable outcomes the five levels of CGIC5 have been ordered with 'Very much or much improved' as the first level and 'Very much or much worse' as the last level. This ordering is the opposite of that required by the lrm function in the rms package but is the same as that required by PROC LOGISTIC and PROC GENMOD in SAS. Since the clm function can fit a variety of cumulative link models the link='logistic' option is used to specify that a proportional odds model or stratified proportional odds model should be fitted.

In models fitted by the clm function the linear predictor has the following form: $\alpha_{sk} - \eta(\underline{z}_i)$. This has $-\eta(\underline{z}_i)$ rather than $+\eta(\underline{z}_i)$ as used in the lrm function in R and the various procedures in SAS. Therefore estimates of the regression coefficients ($\hat{\beta}$s) using the clm function will have the opposite sign. So in this example the estimate of the log odds of a more favourable outcome with Tacrine compared to placebo, stratified by study, is: $\hat{\theta} = -\hat{\beta} = 0.5048$

Model for more favourable outcome stratified by study but without treatment

```
Call:
clm(location = CGIC5 ~ 1, nominal = ~STUDY, data = cgic, weights = COUNT,
    link = "logistic")

No location coefficients

No scale coefficients

Threshold coefficients:
                Estimate Std. Error z value
1|2.(Intercept) -3.5066   0.4143    -8.4633
2|3.(Intercept) -1.1116   0.1614    -6.8860
3|4.(Intercept)  0.9853   0.1566     6.2917
4|5.(Intercept)  3.6939   0.4527     8.1589
1|2.STUDY2       0.1508   0.4906     0.3074
2|3.STUDY2       0.4950   0.1896     2.6116
3|4.STUDY2       0.6269   0.2018     3.1060
4|5.STUDY2       0.1866   0.5643     0.3306
1|2.STUDY3       1.8872   0.4812     3.9217
2|3.STUDY3       0.9626   0.2435     3.9527
3|4.STUDY3       0.4114   0.2766     1.4873
4|5.STUDY3      -0.3180   0.6808    -0.4671
1|2.STUDY4       0.6162   0.4561     1.3509
2|3.STUDY4       0.2092   0.1868     1.1200
3|4.STUDY4       0.0537   0.1842     0.2916
4|5.STUDY4      -0.8393   0.4901    -1.7124
1|2.STUDY5       0.7985   0.6205     1.2869
2|3.STUDY5       0.7061   0.2795     2.5261
3|4.STUDY5       0.8509   0.3605     2.3607
4|5.STUDY5       0.6756   1.1035     0.6122

log-likelihood: -1790.047
AIC: 3620.095
Condition number of Hessian: 937.9847
```

Likelihood ratio test for treatment, stratified by study

```
Likelihood ratio tests of cumulative link models

Response: CGIC5
             Model Resid. df -2logLik   Test    Df LR stat.      Pr(Chi)
1     1 |   | STUDY      1383 3580.095
2 TREAT |   | STUDY      1382 3559.669 1 vs 2     1 20.42567 6.199266e-06
```