**Appendix**

Codes for fitting fixed-effect and mixed-effects logistic regression models in Stata, R and SAS

**Data:**

number of events (death or bleeding) and sample size in each study group (intervention and control) of the 19 sclerotherapy trials (see Table 2 in the paper)

**Variables:**

id= study id (1 to 19)

n= sample size

arm=study groups (0: control, 1: intervention)

dead=number of dead

ndead=n-death

bleed=number patients with bleeding events

nbleed=n-bleed

**Estimation method for mixed-effects logistic regression model**

We used maximum likelihood with adaptive Gauss–Hermite quadrature (AGHQ) based on 7 integration points in all three software packages to estimate the overall intervention effect  and the between-study variance 2 and their variances. AGHQ provides a good approximation to the likelihood function. AGHQ with 7 quadrature points is the default method for melogit in Versions 13 & 14 of Stata. AGHQ with one quadrature point is the default method for glmer in R. Larger numbers of quadrature points provide better approximations to the likelihood function. We used 7 quadrature points because further increases in the number of points did not change the estimates of the parameters. In their analyses users should consider whether more quadrature points are desirable. The default method in SAS glimmix is restricted maximum likelihood (REML) using Taylor-series expansions. Although the REML estimate of 2 is less biased than the ML estimate and is a good alternative when the sole focus is on estimating 2 , we chose to use ML methods because our focus is to estimate both  and 2 and also to be consistent with the method we used in Stata and R.

**Specifying the variable ‘ARM’ as categorical vs. numerical variables**

To obtain the estimate of the overall difference in log (odds) between the two study arms (intervention minus control), it is necessary to place the variable ‘arm’ in the fixed-effect part of the command, but the type of that variable is not the same in the three software packages. In Stata and R the variable ‘arm’ has to be a categorical variable, whereas in SAS it has to be a numerical variable (0 for the control arm and 1 for the intervention arm). With the variable type specification as described the regression coefficient of ‘arm’ is the estimate of the difference in log (odds) between the two study arms. However, ‘arm’ has to be a numerical variable in the random-effect part of the command in all three software packages so that the random effect is across study and only one between-study variance is estimated.

**Stata**

Command: **melogit**

/\* Note: i.id and i.arm denote categorical variables. \*/

**Fixed-effect logistic regression model**

melogit dead i.id i.arm, binomial(n)

 melogit bleed i.id i.arm, binomial(n)

**Mixed-effects logistic regression model**

/\* The default estimation method is maximum likelihood with adaptive Gauss–Hermite quadrature based on 7 integration points. \*/

/\* Note: arm is a numerical variable in the random-effect part of the command. noconstant in the random-effect part of the command is used so that the study effect is fixed \*/

melogit dead i.id i.arm || id: arm, noconstant binomial(n)

melogit bleed i.id i.arm || id: arm, noconstant binomial(n)

**R**

/\* Note: as.factor(id) and as.factor(arm) specifying categorical variables. \*/

**Fixed-effect logistic regression model**

Command: **glm**

Summary (glm(cbind(dead,ndead) ~ as.factor(arm) + as.factor(id), family=binomial))

Summary (glm(cbind(bleed,nbleed)~as.factor(arm) +as.factor(id), family=binomial))

**Mixed-effects logistic regression model**

Command: **glmer**

/\* Note: arm is a numerical variable in the random-effect part of the command. \*/

/\* nAGQ=7 specifies that the estimation method is adaptive Gauss–Hermite quadrature (AGHQ) based on 7 integration points. \*/

Summary (glmer(cbind(dead,ndead) ~ as.factor(arm) +as.factor(id) + (-1+arm|id), family=binomial,nAGQ=7))

Summary (glmer(cbind(bleed,nbleed) ~ as.factor(arm) +as.factor(id) + (-1+arm|id), family=binomial,nAGQ=7))

**SAS**

/\* Note: For both PROC LOGISTIC and PROC GLIMMIX the ‘model’ statement, specifies ARM (1: intervention, 0: control) as a numerical variable, so its regression coefficient indicates the group difference (intervention - control) in log(odds).

If ARM were a categorical variable, the group difference (intervention minus control) in log(odds) would be equal to 2 times the regression coefficient because the design matrix would contain value 1 for the intervention group and –1 for the control group.

ARM is also a numerical variable in the ‘random’ statement for PROC GLIMMIX.

\*/

**Fixed-effect logistic regression model**

**proc** **logistic**;

 class id;

 model dead/n=arm id/risklimits CI;

 **run**;

**proc** **logistic**;

 class id;

 model bleed/n=arm id/risklimits CI;

 **run**;

**Mixed-effects logistic regression model**

/\* method=quad (qpoints=7) specifies that the estimation method is adaptive Gauss–Hermite quadrature (AGHQ) based on 7 integration points. \*/

proc glimmix method=quad (qpoints=7);

 class id;

 model dead/n =arm id /solution cl or;

 random arm/subject=id;

 run;

/\*

proc glimmix method=quad (qpoints=7);

 class id;

 model bleed/n =arm id /solution cl or;

 random arm/subject=id;

 run;