## 1035 Additional file 11 - Final model for assessing heterogeneity of treatment effect

$$\begin{split} \text{logit} \left( P\left(Y_{ki} \geq y \right) \right) &= \alpha + \tau_{yk} + \beta \boldsymbol{X}_{ki} + A_{ki} \left( \delta_{kc} + \gamma_{(ks)c} d_{kis} \right) \\ &\alpha \sim \text{Normal} \left( \mu = 0, \sigma = 0.1 \right) \\ &\tau_{yk} \sim t_{\text{student}} \left( \text{df} = 3, \mu = 0, \sigma = 8 \right) \\ &\beta \sim \text{Normal} \left( \mu = \mathbf{0}, \Sigma = 2.5^2 I_{p \times p} \right) \\ &\delta_{kc} \sim \text{Normal} \left( \mu = \delta_c, \sigma = \eta \right) \\ &\eta \sim t_{\text{student}} \left( \text{df} = 3, \mu = 0, \sigma = 0.25 \right) \\ &\delta_c \sim \text{Normal} \left( \mu = -\Delta, \sigma = 0.1 \right) \\ &-\Delta \sim t_{\text{student}} \left( \text{df} = 3, \mu = 0, \sigma = 2.5 \right) \\ &\gamma_{(ks)c} \sim \text{Normal} \left( \mu = \gamma_{cs}, \sigma = 1 \right) \\ &\gamma_{cs} \sim \text{Normal} \left( \mu = -\Gamma_s, \sigma = 0.25 \right) \\ &-\Gamma_s \sim t_{\text{student}} \left( \text{df} = 3, \mu = 0, \sigma = 1.5 \right) \end{split}$$

1036 The pooled effect of CCP (measured by  $\log \mathsf{OR}$ ) across all RCTs for patients with covariate S=s will be

1037  $\Delta_s = \Delta + \Gamma_s.$ 

1038 The prior distribution of  $\Delta_s$  is less skeptical than  $\Delta_{co}$  in model (7) because we are interested in estimating the

1039 treatment effect in each stratum of S instead of assessing evidence for overall efficacy. A less skeptical prior spreads

out the probability mass, and will not shrink the posterior estimate strongly towards zero; this allows the observed

1041 data to have greater influence over the posterior distribution.

1042 The introduction of parameters  $-\Delta$  and  $-\Gamma_s$  required additional prior distribution assumptions. We found that

using  $t_{\mathsf{student}}$  and Normal prior distributions resulted in similar posterior estimations and divergent transitions. The

1040  $t_{student}$  distribution, however, achieved slightly better model convergence and is less restrictive. Therefore, we opted

1045 for a  $t_{\text{student}}$  prior distribution for both  $-\Delta$  and  $-\Gamma_s$ .