APPENDIX 15: INTERVENTIONS FOR ACUTE DEPRESSION – NETWORK META-ANALYSIS PLOTS AND WINBUGS CODE

Report produced by Sofia Dias and Tony Ades (NICE Clinical Guidelines Technical Support Unit).

Abbreviations

CrI	credible interval
HAMD	Hamilton Rating Scale for Depression
MADRS	Montgomery–Åsberg Depression Rating Scale
N	total number of participants randomised
NA	not applicable
NCCMH	National Collaborating Centre for Mental Health
NICE	National Institute for Health and Care Excellence
NMA	network meta-analysis
SD	standard deviation
SMD	standardised mean difference

1.1 SUMMARY

The methods used in the network meta-analysis (NMA) of pharmacological treatments for acute depression in bipolar disorder are described here. Standard models for NMA with binary outcomes were used for two outcomes: (a) discontinuation; and (b) response given no discontinuation. Information on the log-odds ratio of response in trials reporting on more than one scale was combined and information on the standardised mean difference on different symptoms scales was used to inform the log-odds ratio of response.

Baseline probabilities of discontinuation and response given no discontinuation were calculated based on all trials with a placebo arm reporting these outcomes.

1.2 METHOD

Twenty-seven studies were included in a network meta-analysis (NMA) of the relative effects of 13 pharmacological treatments including placebo for bipolar depression on two outcomes: discontinuation (for any reason) and response, given not discontinued. The 13 treatments compared are presented in Table 15.1 and the 19 available treatment comparisons in Table 15.2 and Figure 15.1.

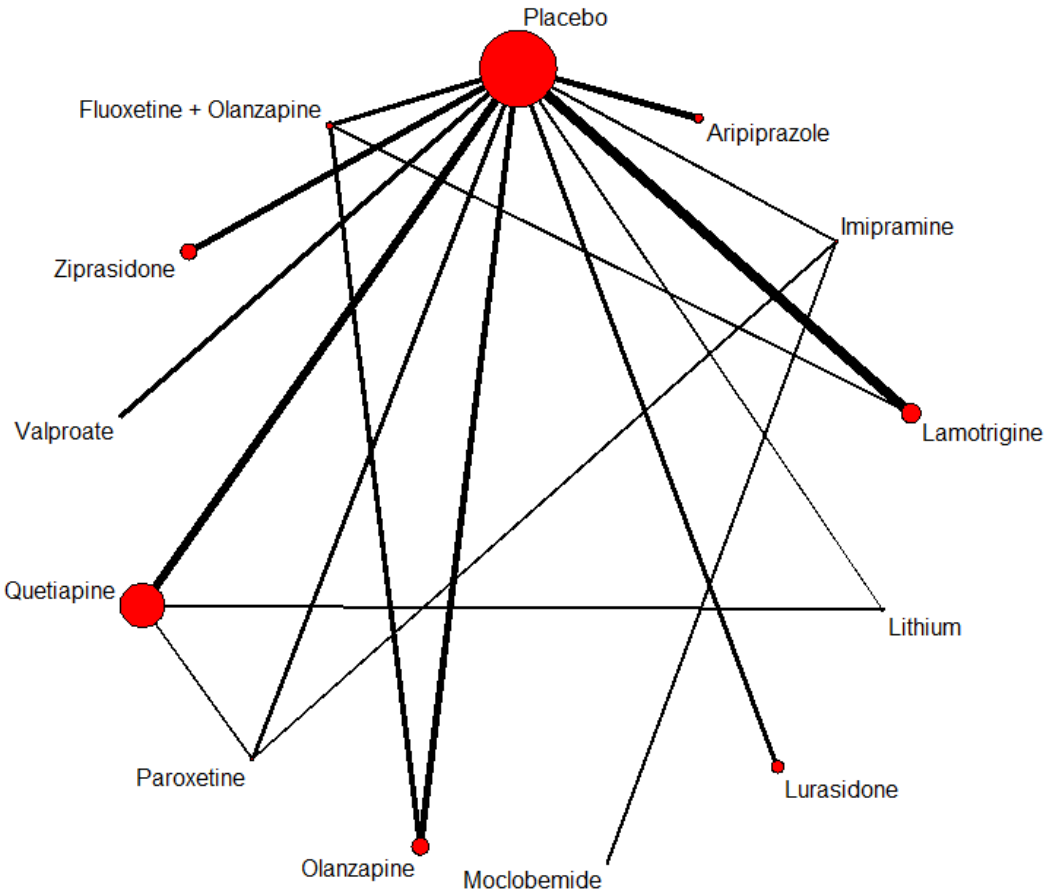
Table 15.1: Treatments compared and number of participants randomised to each treatment

Treatment code	Treatment definition	Total number of participants randomised
1	Placebo (pill)	3,215
2	Aripiprazole	385
3	Imipramine	111
4	Lamotrigine	810
5	Lithium	136
6	Lurasidone	518
7	Moclobemide	81
8	Olanzapine	713
9	Paroxetine	155
10	Quetiapine	1,867
11	Valproate	48
12	Ziprasidone	675
13	Fluoxetine and olanzapine	292

Table 15.2: Number of direct (pairwise) comparisons made and number of participants randomised to each comparison

	t1	t2	Treatments compared		Number of participants randomised
1	1	2	Placebo (pill)	Aripiprazole	772
2	1	3	Placebo (pill)	Imipramine	79
3	1	4	Placebo (pill)	Lamotrigine	1,196
4	1	5	Placebo (pill)	Lithium	269
5	1	6	Placebo (pill)	Lurasidone	853
6	1	8	Placebo (pill)	Olanzapine	1,261
7	1	9	Placebo (pill)	Paroxetine	324
8	1	10	Placebo (pill)	Quetiapine	2,615
9	1	11	Placebo (pill)	Valproate	97
10	1	12	Placebo (pill)	Ziprasidone	1,189
11	1	13	Placebo (pill)	Fluoxetine and olanzapine	464
12	3	7	Imipramine	Moclobemide	156
13	3	9	Imipramine	Paroxetine	69
14	4	13	Lamotrigine	Fluoxetine and olanzapine	410
15	5	10	Lithium	Quetiapine	669
16	8	13	Olanzapine	Fluoxetine and olanzapine	457
17	9	10	Paroxetine	Quetiapine	614

Figure 15.1: Network of treatment comparisons. The thickness of the lines is proportional to the number of trials making that comparison and the size of the nodes is proportional to the number of patients randomised to that treatment



A joint NMA on discontinuation and number of responders given not discontinued was carried out, by subtracting the number of patients who had discontinued from the total number of patients randomised. A separate NMA to estimate relative effects of response out of all randomised patients (that is, not conditional on discontinuation) was also carried out.

Available data on the number of discontinuations and patients randomised is given in Table 15.3. All studies reported the number of patients discontinuing, out of the total number randomised, but only 25 studies reported a useable measure of response on a dichotomous or continuous scale (BRISTOLMYERSSQUIB2006 and BRISTOLMYERSSQUIB2007 did not report response).

Table 15.3: Discontinuation data: disc – number of patients discontinuing for any reason

	Number of trial arms	Treatment											
		Arm 1	Arm 2	Arm 3	Arm 4	Arm 1		Arm 2		Arm 3		Arm 4	
Study		t1	t2	t3	t4	dis c1	N1	dis c2	N2	dis c3	N3	dis c4	N4
BRISTOLMYERSSQUIB 2006	2	1	2	NA	NA	69	188	87	186	NA	NA	NA	NA
BRISTOLMYERSSQUIB 2007	2	1	2	NA	NA	56	188	77	187	NA	NA	NA	NA
QUANTE2010	2	1	2	NA	NA	2	11	3	12	NA	NA	NA	NA
CALABRESE1999	2	1	4	NA	NA	19	66	18	63	NA	NA	NA	NA
CALABRESE2008a	2	1	4	NA	NA	34	103	35	103	NA	NA	NA	NA
CALABRESE2008b	2	1	4	NA	NA	33	124	52	133	NA	NA	NA	NA
CALABRESE2008c	2	1	4	NA	NA	36	110	30	111	NA	NA	NA	NA
CALABRESE2008d	2	1	4	NA	NA	55	128	52	131	NA	NA	NA	NA
VANDERLOOS2009	2	1	4	NA	NA	10	60	12	64	NA	NA	NA	NA
SUNOVION2012a	2	1	6	NA	NA	29	165	40	183	NA	NA	NA	NA
TOHEN2012	2	1	8	NA	NA	49	171	76	343	NA	NA	NA	NA
SUPPES2010	2	1	10	NA	NA	44	140	53	140	NA	NA	NA	NA
DAVIS2005	2	1	11	NA	NA	7	12	6	13	NA	NA	NA	NA
GHAEMI2007	2	1	11	NA	NA	5	9	2	9	NA	NA	NA	NA
MUZINA2011	2	1	11	NA	NA	15	28	13	26	NA	NA	NA	NA
PFIZER2009a	2	1	12	NA	NA	62	196	73	185	NA	NA	NA	NA
SACHS2011	2	1	12	NA	NA	47	150	61	148	NA	NA	NA	NA
SILVERSTONE2001	2	3	7	NA	NA	22	75	27	81	NA	NA	NA	NA
BROWN2006	2	4	13	NA	NA	71	205	68	205	NA	NA	NA	NA
NEMEROFF2001	3	1	3	9	NA	18	43	16	36	10	33	NA	NA
SUNOVION2012b	3	1	6	6	NA	43	170	43	166	46	169	NA	NA
TOHEN2003	3	1	8	13	NA	208	377	191	370	32	87	NA	NA
CALABRESE2005	3	1	10	10	NA	74	181	61	181	83	180	NA	NA
THASE2006	3	1	10	10	NA	58	168	71	172	79	169	NA	NA
PFIZER2009b	3	1	12	12	NA	57	168	63	171	80	171	NA	NA
YOUNG2010	4	1	5	10	10	37	133	34	136	65	265	63	268
MCELROY2010	4	1	9	10	10	50	126	46	122	88	245	88	247

Data on response was reported in different formats. The relative effect of interest was the odds ratio of response, so the following approach was taken to incorporate as much of the available data as possible:

(1) For studies reporting the number of responders on **only one** of the HAMD or MADRS scales, those data were used in the analysis (16 studies: QUANTE2010, VANDERLOOS2009, SUNOVION2012a, TOHEN2012, SUPPES2010, DAVIS2005, MUZINA2011, SILVERSTONE2001, BROWN2006, NEMEROFF2001, SUNOVION2012b, TOHEN2003, CALABRESE2005, THASE2006, YOUNG2010, MCELROY2010). Data in Table 15.4.

(2) For studies reporting the number of responders on **both** the HAMD and MADRS the log-odds ratio of response, given not discontinued, given by each measure was averaged and the standard error of the log-odds ratios was calculated as the average of the standard errors on each scale (five studies: CALABRESE1999, CALABRESE2008c, CALABRESE2008d, PFIZER2009a, PFIZER2009b). PFIZER2009b was a three-arm trial, where two log-odds ratios against the treatment in Arm 1 are calculated for each measure. These within-measure relative effects are correlated and their covariance is equal to the variance of the log-odds in Arm 1[1]. The covariance of the combined log-odds ratio was taken as the average of the covariances on the HAMD and MADRS scales (the resulting covariance matrix was checked to ensure it remained invertible with positive diagonals). Data in Table 15.5.

(3) For studies not reporting the number of responders but reporting the mean and standard deviation (SD) on one of the scales (HAMD or MADRS), the within-study standardised mean difference (SMD) and its variance were calculated according to the Hedges' g formula and used in the analysis (two studies, both only two arms: CALABRESE2008b, SACHS2011). Data in Table 15.6.

(4) For studies not reporting the number of responders but reporting the mean and SD on both the HAMD and MADRS scales, the within-study SMD on each scale and their standard errors were calculated as above, and then averaged. This combined SMD and its variance (the standard error squared) were used in the analysis (two studies, both only two arms: CALABRESE2008a. GHAEMI2007). Data in Table 15.6.

[1] Franchini A, Dias S, Ades AE, Jansen J, Welton N. Accounting for correlation in mixed treatment comparisons with multi-arm trials. Research Synthesis Methods. 2012;3:142-60.

Table 15.4: Data on number of responders available on a single scale

		Treatments				Number of responders			
	Number of trial arms	Arm 1	Arm 2	Arm 3	Arm 4	Arm 1	Arm 2	Arm 3	Arm 4
Study	**na**	**t1**	**t2**	**t3**	**t4**	**r1**	**r2**	**r3**	**r4**
QUANTE2010	2	1	2	NA	NA	9	8	NA	NA
VANDERLOOS2009	2	1	4	NA	NA	19	33	NA	NA
SUNOVION2012a	2	1	6	NA	NA	68	102	NA	NA
TOHEN2012	2	1	8	NA	NA	74	180	NA	NA
SUPPES2010	2	1	10	NA	NA	41	57	NA	NA
DAVIS2005	2	1	11	NA	NA	3	6	NA	NA
MUZINA2011	2	1	11	NA	NA	3	6	NA	NA
SILVERSTONE2001	2	3	7	NA	NA	40	37	NA	NA
BROWN2006	2	4	13	NA	NA	80	95	NA	NA
NEMEROFF2001	3	1	3	9	NA	15	14	15	NA
SUNOVION2012b	3	1	6	6	NA	49	85	83	NA
TOHEN2003	3	1	8	13	NA	108	137	46	NA
CALABRESE2005	3	1	10	10	NA	39	69	57	NA
THASE2006	3	1	10	10	NA	72	93	88	NA
YOUNG2010	4	1	5	10	10	54	64	137	143
MCELROY2010	4	1	9	10	10	40	42	107	107

Table 15.5: Combined log-odds ratios (lor) and variance (Var) data on HAMD and MADRS

		Treatments							
	Number of trial arms	Arm 1	Arm 2	Arm 3	Arm 2 versus Arm 1		Arm 3 versus Arm 1		Co-variance
Study	**na**	**t1**	**t2**	**t3**	**lor2**	**Var2**	**lor3**	**Var3**	**V**
CALABRESE1999	2	1	4	NA	0.85368918	0.188285024	NA	NA	NA
CALABRESE2008c	2	1	4	NA	0.453966334	0.117266514	NA	NA	NA
CALABRESE2008d	2	1	4	NA	0.349976145	0.116197012	NA	NA	NA
PFIZER2009a	2	1	12	NA	0.756870482	0.106686151	NA	NA	NA
PFIZER2009b	3	1	12	12	0.116139986	0.080019649	0.1768027	0.088897426	0.038833533

Table 15.6: SMD and variance data (combined or from single scale)

		Treatments			
		Arm 1	Arm 2	SMD	Variance
Study	**na**	**t1**	**t2**	**Y**	**Var**
CALABRESE2008b	2	1	4	-0.112417114	0.016509948
SACHS2011	2	1	12	-0.021587178	0.013793918
CALABRESE2008a	2	1	4	0.081355061	0.019831094
GHAEMI2007	2	1	11	-0.645198064	0.237027284

Two possible approaches can be adopted to capture the dependency between the two 'competing' outcomes (discontinuation and response): the first is to model response, and discontinuation conditional on no response, as two conditionally independent outcomes. The second is to model discontinuation, and response conditional on no discontinuation. In both cases the correlation between the outcomes is correctly accounted for. They are however **distinctly different models**: to say that differences in response probability are linear on a logit scale is **not** equivalent to saying that differences in response conditional on no discontinuation are linear on a logit scale. Following discussion with the NCCMH technical team, it was agreed that the second option was the one that best captured the clinical situation and provided adequate outputs for the economic model. Network meta-analyses was therefore carried out for (a) discontinuation and (b) response conditional on no discontinuation. A further analysis of response (independently of discontinuation) was also carried out.

Below, the models for baseline (placebo) and relative effects are described.

1.2.1 Baseline probability

To obtain absolute probabilities of discontinuation and response given no discontinuation, it is necessary to obtain a baseline treatment effect for placebo (Treatment 1) to which the relative treatment effects are applied. The baseline probabilities of interest can be estimated by performing a meta-analysis of the placebo arms for the two outcomes of interest: discontinuation and response given not discontinued.

1.2.2 Discontinuation

Twenty-seven studies reporting discontinuation included a placebo arm (Table 15.3). The number of discontinuations in the placebo arm of Trial i (i = 1,…,25) is assumed to follow a binomial likelihood

$$r_{i1}^{D} \sim \text{Binomial}\left(p_{i1}^{D}, n_{i1}\right)$$

where p_{i1}^{D} is the probability of discontinuation and n_{i1} are the total number of patients randomised in Arm 1 (the placebo arm) of Trial i. the probability of discontinuation on the logit scale are modelled

$$\text{logit}(p_{i1}^{D}) = b_i^{D}$$
$$b_i^{D} \sim N(m_D, \tau_D^2)$$

with the trial-specific baselines b_i^D drawn from a distribution of effects. To complete the model, in a Bayesian framework, vague priors were put on the mean, $m_D \sim N(0,100^2)$, and on the variance, $\tau_D \sim \text{Uniform}(0,5)$.

The predictive distributions of the log-odds of discontinuation on placebo in a future trial were approximately normal with posterior means m_D = -0.7025 and standard deviation 0.423, which translates into a baseline probability of discontinuation on placebo of 33% but with a wide 95% credible interval (CrI) from 18% to 53%. These results are used in the relative effects model to generate a baseline A_D~Normal(-0.7025, 0.423^2) on the log-odds scale on which relative effects were added at each iteration, to deliver the posterior summaries on the absolute probability scale for each treatment.

Note that our use of predictive distribution to describe the baseline probabilities greatly increases the degree of uncertainty in this parameter. However, this does no more than reflect the substantial degree of variation between baseline rates in the trials. The higher uncertainty in the baseline has no impact on the **relative** treatment effects, and very little impact on incremental net benefits.

1.2.3 Response given no discontinuation

A shared parameter model[2] was used to incorporate data on the number of responders and on the log-odds of response on placebo from 19 studies including Placebo arm and reporting this outcome (Table 15.4 and Table 15.5).

The number of responders out of the patients who did not discontinue in the placebo arm of Trial *i* (*i*=1,…,14) were modelled in the same way as the number of discontinuations for trials reporting on the number of responders on a single scale, namely the number of responses in the placebo arm of Trial *i* (*i*=1,…,6,8) has a binomial likelihood

$$r_{i1}^R \sim \text{Binomial}\left(p_{i1}^R, N_{i1}\right)$$

where p_{i1}^R is the probability of response given no discontinuation and N_{ik} are the number of patients who did not discontinue in Arm 1 (the placebo arm) of Trial *i*. The conditional probability of response is modelled on the logit scale

$$\text{logit}(p_{i1}^R) = b_i^R$$

For the five studies reporting the number of responders on the HAMD and MADRS scales, the log-odds of response on placebo for each scale and their standard errors were averaged and the combined log-odds were assumed to have a normal likelihood with mean b_i^R and the averaged variance (*i*=15,...,19). The trial-specific

[2] Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 2: a generalised linear modelling framework for pair-wise and network meta-analysis of randomised controlled trials. Technical Support Document. London: NICE; 2011; Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence Synthesis for Decision Making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. Medical Decision Making. 2013;33:607-17.

baselines b_i^R are drawn from a **common** distribution of effects $b_i^R \sim N(m_R, \tau_R^2)$. Vague priors were put on the mean, $m_R \sim N(0,100^2)$, and on the variance, $\tau_R \sim \text{Uniform}(0,5)$.

The predictive distributions of the log-odds of response on placebo, given not discontinued, in a future trial were approximately normal with posterior means 0.16 and standard deviation 0.6064, which translates into a baseline probability of response given no discontinuation on placebo of 54% with 95% CrI 27% to 80%. These results are used in the relative effects model to generate absolute effects, as described above.

1.3 NETWORK META-ANALYSIS MODEL

1.3.1 Discontinuation

A conditional logit model for the two outcomes was used for the probability of discontinuation and the probability of response conditional on no discontinuation. The two outcomes have been modelled jointly as follows. For each Arm *k* of a Trial *i* (*i* = 1,…,27), the number of discontinuations, r_{ik}^D, out of the total number of randomised patients , n_{ik}, are assumed to follow a Binomial distribution

$$r_{ik}^D \sim \text{Binomial}\left(p_{ik}^D, n_{ik}\right)$$

With p_{ik}^D representing the probability of discontinuation in Arm *k* of Trial *i*.
The probabilities of discontinuation are modelled using a random effects network meta-analysis model[3] on the log-odds scale using a logit link such that

$$\text{logit}(p_{ik}^D) = \mu_i^D + \delta_{ik}^D$$

with μ_i^D being given non-informative normal priors and the trial-specific treatment effects of the treatment in Arm *k*, relative to the treatment in Arm 1, δ_{ik}^D, drawn from a common random effects distribution, under the assumption of consistency, with placebo as the reference treatment

$$\delta_{ik}^D \sim N(d_{t_{ik}}^D - d_{t_{i1}}^D, \sigma_D^2)$$

$d_{t_{ik}}^D$ represents the mean effect of the treatment in Arm *k* in Trial *i*, t_{ik}, relative to placebo, and σ_D^2 represents the between-trial variability in treatment effects (heterogeneity), for discontinuation.

[3] Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence Synthesis for Decision Making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. Medical Decision Making. 2013;33:607-17; Lu G, Ades A. Combination of direct and indirect evidence in mixed treatment comparisons. Statistics In Medicine. 2004;23:3105-24.

1.3.2 Response, given no discontinuation

For each Arm *k* of a Trial *i* giving the number of responders (*i*=1,…,19), the likelihood for the number of responders out of those not discontinued is

$$r_{ik}^{R} \sim \text{Binomial}\left(p_{ik}^{R}, N_{ik}\right)$$

where $N_{ik} = n_{ik} - r_{ik}^{D}$ and p_{ik}^{R} is the probability of response conditional on no discontinuation in Arm *k* of Trial *i* which is modelled on the log-odds scale using a logit link such that

$$\text{logit}(p_{ik}^{R}) = \mu_i^{R} + \delta_{ik}^{R}$$

with μ_i^{R} being given non-informative normal priors.

For trials summarised by the pooled log-odds ratio of response given no discontinuation (lor) (Table 15.5), these were modelled as follows

$$lor_{i2} \sim N(\delta_{i2}^{R}, V_{i2}) \quad (1)$$

For trials with more than two treatment arms, the normal likelihood in (1), is replaced with a multivariate normal likelihood for the vector $(lor_{i2}, lor_{i3}, ..., lor_{i,na_i})$ where na_i is the number of treatment arms in Trial *i*, and the correlation between the log-odds ratios calculated in the same multi-arm trial is equal to the variance of the log-odds in Arm 1 of that trial[4].

For trials summarised as SMD (Table 15.6), the SMD, y_{i2}, has been modelled as

$$y_{i2} \sim N(\theta_{i2}^{R}, V_{i2})$$

where θ_{i2}^{R} is the relative treatment effect of the treatment in Arm *k* of Trial *i*, relative to the treatment in Arm 1 on the SMD scale, thus $\theta_{i2}^{R} > 0$ favours the treatment in Arm 1 and $\theta_{i2}^{R} < 0$ favours the treatment in Arm *k*.

The SMD of recovery can be related to a notional lor for response using the formula[5]

$$LOR = -\frac{\pi}{\sqrt{3}} SMD$$

noting the change in sign to retain the interpretation of a positive lor favouring treatment *k*. Therefore the trials summarised as SMD inform the lor of response via the following relationship

$$\theta_{ik}^{R} = \text{-}0.5513 \times \delta_{ik}^{R}$$

---

[4] Franchini A, Dias S, Ades AE, Jansen J, Welton N. Accounting for correlation in mixed treatment comparisons with multi-arm trials. Research Synthesis Methods. 2012;3:142-60.

[5] Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. Statistics In Medicine. 2000;19:3127-31.

This is a shared parameter model[6], where different data points inform the same parameters, δ_{ik}^{R}, through different equations and information on them can be shared within the model. A random effects network meta-analysis model is used to account for between-trial heterogeneity. The trial-specific treatment effects of the treatment in Arm *k*, relative to the treatment in Arm 1, are drawn from a common random effects distribution, under the assumption of consistency:

$$\delta_{ik}^{R} \sim N(d_{1,t_{ik}}^{R} - d_{1,t_{i1}}^{R}, \sigma_{R}^{2})$$

where $d_{1,t_{ik}}^{R}$ represents the mean effect of the treatment in Arm *k* in Trial *i*, t_{ik}, relative to treatment 1 (Placebo), and σ_{R}^{2} represents the between-trial variability in treatment effects on response (heterogeneity). The between-trials standard deviation, σ_{R}, was given a Uniform(0,5) prior. The correlation between the random effects of the trials with more than two arms is taken into account in the analysis[7].

1.3.3 Model properties and assumptions

The model assumes that:

(1) The populations included in all trials are similar and the treatment effects are exchangeable across all patients (for example, the treatment effects are expected to be similar for all included patients and treatments).
(2) The relationship between the lor of 'response given no discontinuation' and the SMD is linear.
(3) The underlying distribution of the SMD is logistic, but can be well approximated by a normal distribution.
(4) The SMD measured on all patients is the same as that measured on non-discontinued patients only.

The model accounts for:

(1) The information provided by multiple measures within the same trial and their correlation.
(2) The correlation between the relative treatment effects in trials with more than two treatments.
(3) The competing nature of the discontinuation and recovery outcomes.

1.3.4 Separate network meta-analysis of response

For each Arm *k* of a Trial *i* giving the number of responders (*i*=1,…,16), the likelihood for the number of responders out of all randomised patients is

$$r_{ik} \sim \text{Binomial}\left(p_{ik}, n_{ik}\right)$$

[6] Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence Synthesis for Decision Making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. Medical Decision Making. 2013;33:607-17.
[7] Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 2: a generalised linear modelling framework for pair-wise and network meta-analysis of randomised controlled trials. Technical Support Document. London: NICE; 2011.

where p_{ik} is the probability of response in Arm *k* of Trial *i* which is modelled on the log-odds scale using a logit link such that

$$\text{logit}(p_{ik}) = \mu_i + \delta_{ik}$$

with μ_i being given non-informative normal priors.

Trials summarised by the pooled log-odds ratio of response (lor) were modelled as follows

$$lor_{i2} \sim N(\delta_{i2}, V_{i2})$$

For trials with more than two treatment arms, the normal likelihood is replaced with a multivariate normal likelihood as described above.

Trials summarised as SMD, y_{i2}, are modelled as $y_{i2} \sim N(\theta_{i2}, V_{i2})$, where θ_{i2} is the relative treatment effect of the treatment in Arm *k* of Trial *i*, relative to the treatment in Arm 1 on the SMD scale and $\theta_{ik} = -0.5513 \times \delta_{ik}$.

The trial-specific treatment effects of the treatment in Arm *k*, relative to the treatment in Arm 1, are drawn from a common random effects distribution and are informed under the assumption of consistency, as before. The correlation between the random effects of the three- and four-arm trials in the network is taken into account in the analysis.

1.3.5 Estimation

Model parameters were estimated using Markov chain Monte Carlo simulation methods implemented in WinBUGS 1.4.3[8]. The first 50,000 iterations were discarded, and 100,000 further iterations were run. In order to test whether prior estimates had an impact on the results, three chains with different initial values were run simultaneously. Convergence was assessed by inspection of the Gelman–Rubin diagnostic plot. Goodness of fit was tested using the posterior mean of the residual deviance, which was compared with the number of data points in the model[9].

The WinBUGS code is provided below.

[8] Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing. 2000;10:325-37; Spiegelhalter DJ. Bayesian methods for cluster randomized trials with continuous responses. Statistics in Medicine. 2001;20:435-52.
[9] Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 2: a generalised linear modelling framework for pair-wise and network meta-analysis of randomised controlled trials. Technical Support Document. London: NICE; 2011.

1.3.6 WinBUGS code – discontinuation and response given no discontinuation

```
# Binomial likelihood, logit link
# Random effects model for multi-arm trials
model{                          # *** PROGRAM STARTS
# model for discontinuation
for(i in 1:ns){                  # LOOP THROUGH STUDIES
   mu[i] ~ dnorm(0,.001)       # vague priors for trial baselines
   w[i,1] <- 0    # adjustment for multi-arm trials is zero for control arm
   delta[i,1] <- 0           # treatment effect is zero for control arm
   for (k in 1:na[i]) {           # LOOP THROUGH ARMS
      disc[i,k] ~ dbin(p[i,k],N[i,k]) # binomial likelihood
      logit(p[i,k]) <- mu[i] + delta[i,k]  # model for linear predictor
      dischat[i,k] <- p[i,k] * N[i,k] # expected value of the numerators
#Deviance contribution
      dev[i,k] <- 2 * (disc[i,k] * (log(disc[i,k])-log(dischat[i,k]))
    + (N[i,k]-disc[i,k]) * (log(N[i,k]-disc[i,k]) - log(N[i,k]-dischat[i,k])))
     }
#  summed residual deviance contribution for this trial
   resdev[i] <- sum(dev[i,1:na[i]])
   for (k in 2:na[i]) {           # LOOP THROUGH ARMS
# trial-specific LOR distributions
      delta[i,k] ~ dnorm(md[i,k],taud[i,k])
# mean of LOR distributions (with multi-arm trial correction)
      md[i,k] <-  d[t[i,k]] - d[t[i,1]] + sw[i,k]
# precision of LOR distributions (with multi-arm trial correction)
      taud[i,k] <- tau *2*(k-1)/k
# adjustment for multi-arm RCTs
      w[i,k] <- (delta[i,k] - d[t[i,k]] + d[t[i,1]])
# cumulative adjustment for multi-arm trials
      sw[i,k] <- sum(w[i,1:k-1])/(k-1)
     }
  }
totresdev <- sum(resdev[])      # Total Residual Deviance for disc
d[1]<-0      # treatment effect is zero for reference treatment
# vague priors for treatment effects
for (k in 2:nt){  d[k] ~ dnorm(0,.0001) }
sd ~ dunif(0,5)     # vague prior for between-trial SD
tau <- pow(sd,-2)   # between-trial precision = (1/between-trial variance)
# ranking discontinuation on relative scale
for (k in 1:nt) {
   rk[k]<- rank(d[],k)          # smallest is best
   best[k] <- equals(rk[k],1)    # rank=1 is best
#calculate probability that treat k is h-th best for discontinuation
   for (h in 1:nt) { prob[h,k] <- equals(rk[k],h) }
  }
```

```
# pairwise ORs of discontinuation
for (c in 1:(nt-1)) {
   for (k in (c+1):nt)  {
      OR[c,k] <- exp(d[k] - d[c])
      LOR[c,k]<-(d[k]-d[c])
    }
 }
# Provide estimates of probability of discontinuation T[k]
# Given a Mean Effect, meanA, for 'standard' treatment A,
# with precision (1/variance) precA
A ~ dnorm(meanA,precA)
for (k in 1:nt) { logit(T[k]) <- A + d[k]  }
#
# model for Response | no discontinuation
#
# Arm-based counts, conditional Binomial, logit link
for(i in 2:nsRA){  # LOOP THROUGH STUDIES WITH ONE SCALE ONLY (except
first one)
   for (k in 1:na[IDRA[i]]) {   # LOOP THROUGH ARMS (given in disc data)
      n[i,k] <- N[IDRA[i],k] - disc[IDRA[i],k] # conditional denominators
      r[i,k] ~ dbin(pR[i,k],n[i,k]) # binomial likelihood
    }
}
for (k in 1:2) {   # Correct first study for 9/9 events
   n[1,k] <- N[IDRA[1],k] - disc[IDRA[1],k] +1 # conditional denominators
   r1[1,k] <- r[1,k] + 0.5
   r1[1,k] ~ dbin(pR[1,k],n[1,k]) # binomial likelihood
}
for(i in 1:nsRA){              # LOOP THROUGH STUDIES WITH ONE SCALE ONLY
   wRA[i,1] <- 0    # adjustment for multi-arm trials is zero for control arm
   deltaRA[i,1] <- 0           # treatment effect is zero for control arm
   muRA[i] ~ dnorm(0,.0001)
   for (k in 1:na[IDRA[i]]) {   # LOOP THROUGH ARMS (given in disc data)
      logit(pR[i,k]) <- muRA[i] + deltaRA[i,k]  # model for linear predictor
      rhat[i,k] <- pR[i,k] * n[i,k] # expected value of the numerators
#Deviance contribution
      devRA[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k]))
    +  (n[i,k]-r[i,k]) * (log(n[i,k]-r[i,k]) - log(n[i,k]-rhat[i,k])))
     }
#  summed residual deviance contribution for this trial
   resdevRA[i] <- sum(devRA[i,1:na[IDRA[i]]])
   for (k in 2:na[IDRA[i]]) {   # LOOP THROUGH ARMS
# trial-specific LOR distributions
      deltaRA[i,k] ~ dnorm(mdRA[i,k],taudRA[i,k])
# mean of LOR distributions (with multi-arm trial correction)
      mdRA[i,k] <-  dR[t[IDRA[i],k]] - dR[t[IDRA[i],1]] + swRA[i,k]
```

```
# precision of LOR distributions (with multi-arm trial correction)
      taudRA[i,k] <- tauR *2*(k-1)/k
# adjustment for multi-arm RCTs
      wRA[i,k] <- (deltaRA[i,k] - dR[t[IDRA[i],k]] + dR[t[IDRA[i],1]])
# cumulative adjustment for multi-arm trials
      swRA[i,k] <- sum(wRA[i,1:k-1])/(k-1)
    }
  }
totresdevRA <- sum(resdevRA[]) # Total Residual Deviance for response arm data
#
# Pooled LOR, Normal likelihood (treatment differences), identity link
for(i in 1:nsL2) {              # LOOP THROUGH 2-ARM STUDIES with LOR
  lor[i,2] ~ dnorm(deltaL[i,2],precL[i,2]) # likelihood for 2-arm trials
# Deviance contribution for trial i
  resdevL[i] <- (lor[i,2]-deltaL[i,2])*(lor[i,2]-deltaL[i,2])*precL[i,2]
 }
for(i in (nsL2+1):(nsL2+nsL3)) {   # LOOP THROUGH THREE-ARM STUDIES with
LOR
  for (k in 1:(na[IDL[i]]-1)) {    # set variance-covariance matrix
    for (j in 1:(na[IDL[i]]-1)) {
      Sigma[i,j,k] <- V[i]*(1-equals(j,k)) + Var[i,k+1]*equals(j,k)
     }
    }
# Precision matrix
  Omega[i,1:(na[IDL[i]]-1),1:(na[IDL[i]]-1)] <- inverse(Sigma[i,,])
# multivariate normal likelihood for 3-arm trials
  lor[i,2:na[IDL[i]]] ~ dmnorm(deltaL[i,2:na[IDL[i]]],Omega[i,1:(na[IDL[i]]-
1),1:(na[IDL[i]]-1)])
# Deviance contribution for trial i
  for (k in 1:(na[IDL[i]]-1)){  # multiply vector & matrix
    lordiff[i,k] <- lor[i,(k+1)] - deltaL[i,(k+1)]
   z[i,k]<- inprod2(Omega[i,k,1:(na[IDL[i]]-1)], lordiff[i,1:(na[IDL[i]]-1)])
                }
  resdevL[i]<- inprod2(lordiff[i,1:(na[IDL[i]]-1)], z[i,1:(na[IDL[i]]-1)])
 }
for(i in 1:(nsL2+nsL3)){                #   LOOP THROUGH ALL STUDIES
  wL[i,1] <- 0    # adjustment for multi-arm trials is zero for control arm
  deltaL[i,1] <- 0          # treatment effect is zero for control arm
  for (k in 2:na[IDL[i]]) {       #  LOOP THROUGH ARMS
    precL[i,k] <- 1/Var[i,k]      # set precisions
# trial-specific LOR distributions
    deltaL[i,k] ~ dnorm(mdL[i,k],taudL[i,k])
# mean of random effects distributions, with multi-arm trial correction
    mdL[i,k] <-  dR[t[IDL[i],k]] - dR[t[IDL[i],1]] + swL[i,k]
# precision of random effects distributions (with multi-arm trial correction)
    taudL[i,k] <- tauR *2*(k-1)/k
```

```
# adjustment, multi-arm RCTs
      wL[i,k] <- (deltaL[i,k] - dR[t[IDL[i],k]] - dR[t[IDL[i],1]])
# cumulative adjustment for multi-arm trials
      swL[i,k] <- sum(wL[i,1:k-1])/(k-1)
    }
  }
totresdevL <- sum(resdevL[])    # Total Residual Deviance for LOR data
#
# SMD HAMD/MADRS or pooled, Normal likelihood, identity link
for(i in 1:nsD) {              # LOOP THROUGH 2-ARM STUDIES with SMD
   y[i,2] ~ dnorm(theta[i,2],prec[i,2]) # likelihood for 2-arm trials
# Deviance contribution for trial i
   resdevD[i] <- (y[i,2]-theta[i,2])*(y[i,2]-theta[i,2])*prec[i,2]
  }
for(i in 1:nsD){                  #   LOOP THROUGH ALL STUDIES
   wD[i,1] <- 0    # adjustment for multi-arm trials is zero for control arm
   deltaD[i,1] <- 0           # treatment effect is zero for control arm
   for (k in 2:na[IDD[i]]) {       #  LOOP THROUGH ARMS
      theta[i,k] <- -0.5513 * deltaD[i,k] # convert SMD to LOR (change sign)
      prec[i,k] <- 1/Vary[i,k]      # set precisions
# trial-specific LOR distributions
      deltaD[i,k] ~ dnorm(mdD[i,k],taudD[i,k])
# mean of random effects distributions, with multi-arm trial correction
      mdD[i,k] <-  dR[t[IDD[i],k]] - dR[t[IDD[i],1]] + swD[i,k]
# precision of random effects distributions (with multi-arm trial correction)
      taudD[i,k] <- tauR *2*(k-1)/k
# adjustment, multi-arm RCTs
      wD[i,k] <- (deltaD[i,k] - dR[t[IDD[i],k]] - dR[t[IDD[i],1]])
# cumulative adjustment for multi-arm trials
      swD[i,k] <- sum(wD[i,1:k-1])/(k-1)
    }
  }
totresdevD <- sum(resdevD[])    # Total Residual Deviance for SMD data
#
# Shared parameters for Response | no discontinuation
dR[1]<-0      # treatment effect is zero for reference treatment
# vague priors for treatment effects
for (k in 2:nt){  dR[k] ~ dnorm(0,.0001) }
sdR ~ dunif(0,5)     # vague prior for between-trial SD
tauR <- pow(sdR,-2)   # between-trial precision = (1/between-trial variance)
# ranking on relative scale for response
for (k in 1:nt) {
   rkR[k]<- (nt+1)-rank(dR[],k)    # larger is best
   bestR[k] <- equals(rkR[k],1)    # rank=1 is best
#calculate probability that treat k is h-th best for response
   for (h in 1:nt) { probR[h,k] <- equals(rkR[k],h) }
```

```
  }
# pairwise ORs for response
for (c in 1:(nt-1)) {
   for (k in (c+1):nt)  {
      ORR[c,k] <- exp(dR[k] - dR[c])
      LORR[c,k]<-(dR[k]-dR[c])
    }
 }
# Provide estimates of probability of response TR[k]
# Given a Mean Effect, meanAR, for 'standard' treatment A,
# with precision (1/variance) precAR
AR ~ dnorm(meanAR,precAR)
for (k in 1:nt) { logit(TR[k]) <- AR + dR[k]  }
}                              # *** PROGRAM ENDS
```

1.3.7 WinBUGS code – response in all randomised participants

```
# Random effects model for multi-arm trials
model{                         # *** PROGRAM STARTS
# model for Response
# Arm-based counts, Binomial, logit link
for(i in 1:nsRA){  # LOOP THROUGH STUDIES WITH ONE SCALE ONLY
   for (k in 1:na[IDRA[i]]) {   # LOOP THROUGH ARMS (given in disc data)
      n[i,k] <- N[IDRA[i],k]   # ITT denominators
      r[i,k] ~ dbin(pR[i,k],n[i,k]) # binomial likelihood
    }
}
for(i in 1:nsRA){              # LOOP THROUGH STUDIES WITH ONE SCALE ONLY
   wRA[i,1] <- 0    # adjustment for multi-arm trials is zero for control arm
   deltaRA[i,1] <- 0           # treatment effect is zero for control arm
   muRA[i] ~ dnorm(0,.0001)
   for (k in 1:na[IDRA[i]]) {   # LOOP THROUGH ARMS (given in disc data)
      logit(pR[i,k]) <- muRA[i] + deltaRA[i,k]  # model for linear predictor
      rhat[i,k] <- pR[i,k] * n[i,k] # expected value of the numerators
#Deviance contribution
      devRA[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k]))
    +  (n[i,k]-r[i,k]) * (log(n[i,k]-r[i,k]) - log(n[i,k]-rhat[i,k])))
     }
#  summed residual deviance contribution for this trial
   resdevRA[i] <- sum(devRA[i,1:na[IDRA[i]]])
   for (k in 2:na[IDRA[i]]) {   # LOOP THROUGH ARMS
# trial-specific LOR distributions
      deltaRA[i,k] ~ dnorm(mdRA[i,k],taudRA[i,k])
# mean of LOR distributions (with multi-arm trial correction)
      mdRA[i,k] <-  dR[t[IDRA[i],k]] - dR[t[IDRA[i],1]] + swRA[i,k]
# precision of LOR distributions (with multi-arm trial correction)
```

```
      taudRA[i,k] <- tauR *2*(k-1)/k
# adjustment for multi-arm RCTs
      wRA[i,k] <- (deltaRA[i,k] - dR[t[IDRA[i],k]] + dR[t[IDRA[i],1]])
# cumulative adjustment for multi-arm trials
      swRA[i,k] <- sum(wRA[i,1:k-1])/(k-1)
    }
  }
totresdevRA <- sum(resdevRA[]) # Total Residual Deviance for response arm data
#
# Pooled LOR, Normal likelihood (treatment differences), identity link
for(i in 1:nsL2) {              # LOOP THROUGH 2-ARM STUDIES with LOR
  lor[i,2] ~ dnorm(deltaL[i,2],precL[i,2]) # likelihood for 2-arm trials
# Deviance contribution for trial i
  resdevL[i] <- (lor[i,2]-deltaL[i,2])*(lor[i,2]-deltaL[i,2])*precL[i,2]
  }
for(i in (nsL2+1):(nsL2+nsL3)) {   # LOOP THROUGH THREE-ARM STUDIES with
LOR
  for (k in 1:(na[IDL[i]]-1)) {    # set variance-covariance matrix
    for (j in 1:(na[IDL[i]]-1)) {
      Sigma[i,j,k] <- V[i]*(1-equals(j,k)) + Var[i,k+1]*equals(j,k)
     }
    }
# Precision matrix
  Omega[i,1:(na[IDL[i]]-1),1:(na[IDL[i]]-1)] <- inverse(Sigma[i,,])
# multivariate normal likelihood for 3-arm trials
  lor[i,2:na[IDL[i]]] ~ dmnorm(deltaL[i,2:na[IDL[i]]],Omega[i,1:(na[IDL[i]]-
1),1:(na[IDL[i]]-1)])
# Deviance contribution for trial i
  for (k in 1:(na[IDL[i]]-1)){  # multiply vector & matrix
    lordiff[i,k] <- lor[i,(k+1)] - deltaL[i,(k+1)]
   z[i,k]<- inprod2(Omega[i,k,1:(na[IDL[i]]-1)], lordiff[i,1:(na[IDL[i]]-1)])
                  }
  resdevL[i]<- inprod2(lordiff[i,1:(na[IDL[i]]-1)], z[i,1:(na[IDL[i]]-1)])
  }
for(i in 1:(nsL2+nsL3)){                  #   LOOP THROUGH ALL STUDIES
  wL[i,1] <- 0    # adjustment for multi-arm trials is zero for control arm
  deltaL[i,1] <- 0          # treatment effect is zero for control arm
  for (k in 2:na[IDL[i]]) {       #  LOOP THROUGH ARMS
    precL[i,k] <- 1/Var[i,k]      # set precisions
# trial-specific LOR distributions
      deltaL[i,k] ~ dnorm(mdL[i,k],taudL[i,k])
# mean of random effects distributions, with multi-arm trial correction
      mdL[i,k] <-  dR[t[IDL[i],k]] - dR[t[IDL[i],1]] + swL[i,k]
# precision of random effects distributions (with multi-arm trial correction)
      taudL[i,k] <- tauR *2*(k-1)/k
# adjustment, multi-arm RCTs
```

```
      wL[i,k] <- (deltaL[i,k] - dR[t[IDL[i],k]] - dR[t[IDL[i],1]])
# cumulative adjustment for multi-arm trials
      swL[i,k] <- sum(wL[i,1:k-1])/(k-1)
    }
  }
totresdevL <- sum(resdevL[])    # Total Residual Deviance for LOR data
#
# SMD HAMD/MADRS or pooled, Normal likelihood, identity link
for(i in 1:nsD) {              # LOOP THROUGH 2-ARM STUDIES with SMD
   y[i,2] ~ dnorm(theta[i,2],prec[i,2]) # likelihood for 2-arm trials
# Deviance contribution for trial i
   resdevD[i] <- (y[i,2]-theta[i,2])*(y[i,2]-theta[i,2])*prec[i,2]
  }
for(i in 1:nsD){                  #   LOOP THROUGH ALL STUDIES
   wD[i,1] <- 0    # adjustment for multi-arm trials is zero for control arm
   deltaD[i,1] <- 0           # treatment effect is zero for control arm
   for (k in 2:na[IDD[i]]) {       #  LOOP THROUGH ARMS
      theta[i,k] <- -0.5513 * deltaD[i,k] # convert SMD to LOR (change sign)
      prec[i,k] <- 1/Vary[i,k]      # set precisions
# trial-specific LOR distributions
      deltaD[i,k] ~ dnorm(mdD[i,k],taudD[i,k])
# mean of random effects distributions, with multi-arm trial correction
      mdD[i,k] <-  dR[t[IDD[i],k]] - dR[t[IDD[i],1]] + swD[i,k]
# precision of random effects distributions (with multi-arm trial correction)
      taudD[i,k] <- tauR *2*(k-1)/k
# adjustment, multi-arm RCTs
      wD[i,k] <- (deltaD[i,k] - dR[t[IDD[i],k]] - dR[t[IDD[i],1]])
# cumulative adjustment for multi-arm trials
      swD[i,k] <- sum(wD[i,1:k-1])/(k-1)
    }
  }
totresdevD <- sum(resdevD[])    # Total Residual Deviance for SMD data
#
# Shared parameters for Response
dR[1]<-0      # treatment effect is zero for reference treatment
# vague priors for treatment effects
for (k in 2:nt){  dR[k] ~ dnorm(0,.0001) }
sdR ~ dunif(0,5)     # vague prior for between-trial SD
tauR <- pow(sdR,-2)   # between-trial precision = (1/between-trial variance)
# ranking on relative scale for response
for (k in 1:nt) {
   rkR[k]<- (nt+1)-rank(dR[],k)    # larger is best
   bestR[k] <- equals(rkR[k],1)    # rank=1 is best
#calculate probability that treat k is h-th best for response
   for (h in 1:nt) { probR[h,k] <- equals(rkR[k],h) }
  }
```

```
# pairwise ORs for response
for (c in 1:(nt-1)) {
   for (k in (c+1):nt)  {
      ORR[c,k] <- exp(dR[k] - dR[c])
      LORR[c,k]<-(dR[k]-dR[c])
    }
 }
# Provide estimates of probability of response T[k]
# Given a Mean Effect, meanA, for 'standard' treatment A,
# with precision (1/variance) precA
# AR ~ dnorm(meanAR,precAR)
# for (k in 1:nt) { logit(TR[k]) <- AR + dR[k]  }
}                              # *** PROGRAM ENDS
```