Appendix Hoofdstuk 7b Wetenschappelijke onderbouwing farmacotherapie bij bipolaire depressie

7.3.2.1 Clinical review protocol (pharmacological and nutritional interventions for acute episodes of bipolar depression)

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 6 (a complete list of review questions and protocols can be found in Appendix 8; further information about the search strategy can be found in Appendix 9).

Table 8: Clinical review protocol summary for the review of pharmacological and nutritional interventions for acute episodes of bipolar depression

Topic	Interventions
Review question	RQ2.2: For adults with bipolar disorder, what are the relative benefits and harms of pharmacological and nutritional interventions for acute episodes of acute bipolar depression? What amendments, if any, need to be made for (i) particular cultural or minority ethnic groups, (ii) gender, (iii) adults (18 to 64) and older adults (65+)?
Objectives	To estimate the efficacy of interventions to treat acute episodes of bipolar depression.
Criteria for considering studies for the review	
Intervention	All licensed oral medications (and their combinations). Nutritional interventions will be analysed separately.
Comparator	Placebo Other interventions
Types of participants	Adults (18+) with bipolar disorder who are experiencing an acute episodes of bipolar depression. Special consideration will be given to the groups above.
Outcomes	Response (50% reduction in symptoms) Discontinuation (due to side effect, other)
Time	The main analysis will include outcomes at the end of the acute treatment phase.

Study design	Randomised controlled trials (RCTs) and cluster RCTs with a parallel group design in which providers and participants were blind to treatment. Quasi-RCTs, such as trials in which allocation is determined by alternation or date of birth, and single-blind studies, will be excluded.
Dosage	Fixed or flexible doses within the therapeutic range (BNF recommended).
Minimum sample size	To be included in a network meta-analysis, drugs must have been evaluated in at least 20 participants.
Study setting	Primary, secondary, tertiary, health and social care
Note. BNF = British National Formulary.	

7.3.2.2 Studies considered

Twenty-seven RCTs (N = 9,006) published between 1999 and 2012 compared eligible interventions and reported outcomes that could be used for network meta-analysis: BRISTOLMYERSSQUIB2006 (Bristol-Myers Squibb, 2006; Thase et al., 2008), BRISTOLMYERSSQUIB2007 (Bristol-Myers Squibb, 2007; Thase et al., 2008), BROWN2006 (Brown et al., 2009; Brown et al., 2006; Nierenberg, 2007), CALABRESE1999 (Bowden, 1999; Calabrese et al., 1999; GlaxoSmithKline, 2005a; GlaxoSmithKline, 2005d; McElroy et al., 2004; Preston et al., 2004; Rudd, 1998), CALABRESE2005 (Calabrese et al., 2005; Cookson et al., 2007; Endicott et al., 2008; Endicott et al., 2007; Hirschfeld et al., 2006; Tohen et al., 2013; Weisler et al., 2008), CALABRESE2008a (Calabrese et al., 2008; Geddes et al., 2009; GlaxoSmithKline, 2005b; GlaxoSmithKline, 2005e; Goldsmith et al., 2004), CALABRESE2008b (Calabrese et al., 2008; Geddes et al., 2009; Goldsmith et al., 2004), CALABRESE2008c (Calabrese et al., 2008; Geddes et al., 2009), CALABRESE2008d (Calabrese et al., 2008; Geddes et al., 2009), DAVIS2005 (Davis et al., 2005), GHAEMI2007 (Ghaemi, 2005; Ghaemi et al., 2007), MCELROY2010 (McElroy et al., 2010; Young et al., 2012; Young, 2008), MUZINA2011 (Muzina, 2008; Muzina et al., 2011), NEMEROFF2001 (GlaxoSmithKline, 2005c; Nemeroff et al., 2001), PFIZER2009a (Gao et al., 2013; Lombardo et al., 2012; Pfizer, 2009a), PFIZER2009b (Gao et al., 2013; Lombardo et al., 2012; Pfizer, 2009b), QUANTE2010 (Quante et al., 2010), SACHS2011 (Sachs et al., 2011), SILVERSTONE2001 (Silverstone, 1997; Silverstone, 2001), SUNOVION2012a (Citrome et al., 2013; Ketter et al., 2012; Sunovion, 2012b), SUNOVION2012b (Citrome et al., 2013; Ketter et al., 2012; Sunovion, 2012a), SUPPES2010 (Suppes et al., 2010), THASE2006 (Endicott et al., 2008; Goodwin, 2007; Thase, 2007; Thase et al., 2006; Tohen et al., 2013; Weisler et al., 2008), TOHEN2003 (Corya et al., 2006; Dube et al., 2007; Shi et al., 2004; Tohen et al., 2007; Tohen et al., 2003; Vieta et al., 2009), TOHEN2012 (Katagiri et al., 2013; Tohen et al., 2012), VANDERLOOS2009 (van der Loos et al., 2010; van der Loos et al., 2011; van der Loos et al., 2009),

YOUNG2010 (Grunze, 2010; Young et al., 2010). Six of these were unpublished (BRISTOLMYERSSQUIB2006, BRISTOLMYERSSQUIB2007, PFIZER2009a, PFIZER2009b, SUNOVION2012a, SUNOVION2012b). Studies included in the network meta-analysis were analysed by comparing discontinuation (for any reason) and response, given not discontinued.

A joint network meta-analysis 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 network meta-analysis to estimate relative effects of response out of all randomised patients (that is, not conditional on discontinuation) was also carried out.

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).

Data on response were 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:

For studies reporting the number of responders on only one of the HAMD or MADRS scales, those data were used in the analysis.

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

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.

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.

One additional three-arm study (N = 174; POST2006) was a comparison of three drugs that could not be connected to the network. Therefore, the pairwise comparisons are reported separately below.

An additional 26 studies were excluded; eight were open-label studies: AMSTERDAM2009 (Amsterdam & Shults, 2009), ASTRAZENECA2012a

(Astrazeneca, 2012a), ASTRAZENECA2012b (Astrazeneca, 2012b), NIERENBERG2006 (Nierenberg et al., 2006), NOLEN2007 (Nolen et al., 2007), TAMAYO2009 (Tamayo et al., 2009), WANG2010 (Wang et al., 2010), YONGNING2005 (Yong Ning & Hui, 2005); six trials were of medications neither routinely used nor licensed for the treatment of mental health problems: CHENGAPPA2000 (Chengappa et al., 2000), DENICOFF2005 (Denicoff et al., 2005) DIAZGRANADOS2010 (Diazgranados et al., 2010), FUREY2013 (Furey & Zarate, 2013), STAMM2011 (Stamm, 2011), SZUBA2005 (Szuba et al., 2005), WATSON2012 (Watson et al., 2012), YOUNG2004 (Young et al., 2004); and four trials included people who did not have bipolar disorder: FIEVE1968 (Fieve et al., 1968), KESSELL1975 (Kessell & Holt, 1975), SMITH1978 (Smith et al., 1978), SPEER2009 (Speer et al., 2009). Three studies were excluded because did not include a sufficient number of participants to be included; one was a study of pramipexole as a second-line intervention: GOLDBERG2004 (Goldberg et al., 2004); one was a study of pramipexole: ZARATE2004B (Zarate et al., 2004); and one was a study of risperidone and paroxetine: SHELTON2004 (Shelton & Stahl, 2004). One study of tranylcypromine was excluded because it did not report response on an accepted measure: HIMMELHOCH1991 (Himmelhoch et al., 1991). Two studies were excluded because they did not report usable outcomes; one compared olanzapine and fluoxetine alone or in combination: AMSTERDAM2005a (Amsterdam & Shults, 2005); one compared valproate with lithium: OQUENDO2011 (Oquendo et al., 2011). One study of eicosapentaenoic acid was excluded because there were only six participants in each group: OSHER2005 (Osher et al., 2005). One was excluded because participants were not acutely depressed: FRANGOU2006 (Frangou et al., 2006). Results could not be obtained for five studies: AHUJA2011 (Ahuja, 2011), COLOMBO2000 (Colombo et al., 2000), FOREST2010 (Forest, 2010), GAO2008 (Gao et al., 2008), MCELROY2013 (McElroy et al., 2013); although they have published several papers about the drug, the manufacturer of cariprazine has not reported the results of clinical trials, and they refused requests for data.

Further information about both included and excluded studies can be found in Appendix 15 and Appendix 32.

7.3.2.3 Network meta-analysis of pharmacological interventions for acute episodes of bipolar depression

Trials included in the network meta-analysis included between 19 and 833 participants at baseline (median 298). Where known, participants were on average (median of means) aged 40 years and about 58% of them were female. Fourteen trials included only participants with bipolar I disorder; one trial included only participants with bipolar II disorder (CALABRESE2008c), and only 37% of participants in another had bipolar II disorder (MUZINA2011).

Studies of medication alone or as an addition to another treatment were included. All participants were taking a mood stabiliser in six studies (QUANTE2010, SACHS2011, NEMEROFF2001, VANDERLOOS2009,

SUNOVION2012a, SUNOVION2012b). Twelve studies reported that participants were not taking mood stabilisers at baseline (BRISTOLMYERSSQUIB2006, BRISTOLMYERSSQUIB2007, CALABRESE1999, CALABRESE2005, CALABRESE2008a, CALABRESE2008b, CALABRESE2008c, CALABRESE2008d, DAVIS2005, GHAEMI2007, MCELROY2010, MUZINA2011, PFIZER2009a, PFIZER2009b, SUPPES2010, THASE2006, TOHEN2003, YOUNG2010), though participants in some of these studies could be taking other medications including anxiolytics or hypnotics. Nine studies included a mix of participants taking or not taking mood stabilisers, or did not report their use.

Quality of the evidence

To rate the quality of evidence, guidelines may use GRADE profiles for critical outcomes. However, GRADE has not yet been adapted for use in network meta-analyses. To evaluate the quality of the evidence from the network meta-analysis, information about the factors that would normally be included in a GRADE profile will be reported (that is, risk of bias, publication bias, imprecision, inconsistency and indirectness).

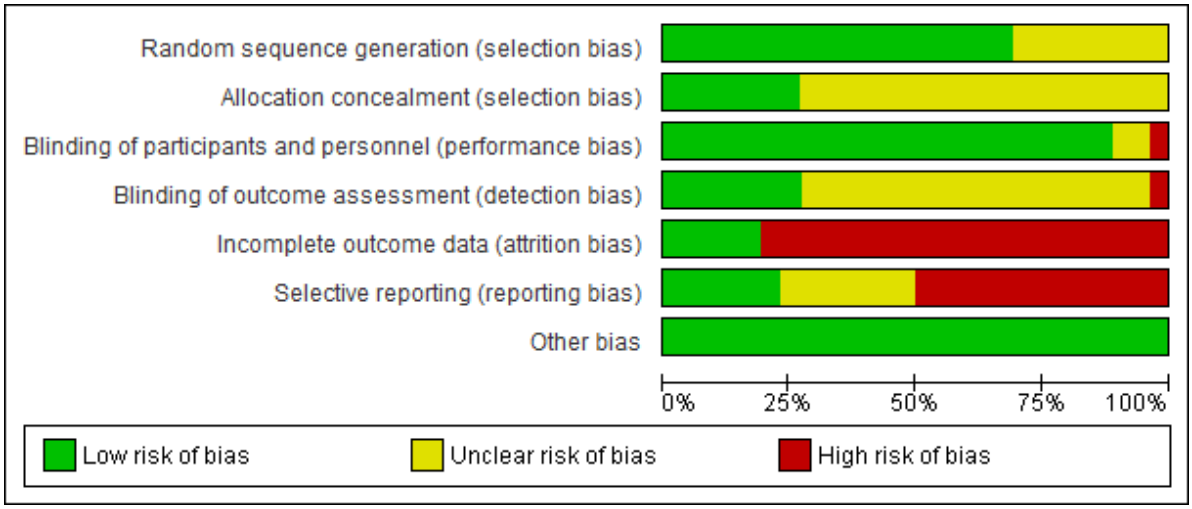
Risk of bias

All included trials were assessed for risk of bias (Appendix 13). Of those in the network meta-analysis, 21 were at low risk for sequence generation and nine of these were at low risk of bias for allocation concealment. Allocation concealment was unclear in 18 trials. All trials were double-blind and were rated as low risk of bias for participant and provider blinding, although effects of medication, including side effect, may make it difficult to maintain participant and provider blinding, particularly at higher doses. Assessor blinding was considered separately for all trials; seven were at low risk of bias and assessors were aware of treatment conditions in one trial. For incomplete outcome data, response was analysed assuming that participants who discontinued treatment did not respond. Because of the high rate of missing data and/or the handling of missing data, continuous outcomes were at high risk of bias in 22 trials.

Selective outcome reporting and publication bias

Several methods were employed to minimise risk of selective outcome reporting and publication bias. The NCCMH review team wrote to all authors to request trial registrations and unpublished outcomes, and all authors of included trials, all stakeholders, and pharmaceutical manufacturers were asked to provide unpublished trials. Nonetheless, only six were at low risk of selective outcome reporting bias, the remaining 14 and seven were at unclear and high risk of bias.

Figure 2: Risk of bias summary



Inconsistency

Inconsistency was assessed by fitting an unrelated mean effects model (Dias et al., 2012) and comparing the fit of this model to the fit of the full network meta-analysis model using the residual deviance (Dias et al., 2012). The posterior mean of the residual deviance for discontinuation was 63.5, very close to the respective 64 data points of the model; the posterior mean of the total residual deviance for response was 58.44, moderately high compared with the respective 51 data points. This finding may be attributable to one study (THASE2006) that did not fit the model well regarding response. Only one loop in the network had the potential for inconsistency, and there was no evidence of inconsistency for response and for discontinuation.

Indirectness

All evidence in the network meta-analysis is direct insofar as it relates to the population, interventions and outcomes of interest.

Effects of interventions

In the network meta-analysis, all interventions except aripiprazole ranked higher than placebo for response given no discontinuation, but only six were statistically superior to placebo (lurasidone, valproate, quetiapine, the combination of fluoxetine and olanzapine, olanzapine alone, and lamotrigine) (see Table 7). Quetiapine and lurasidone were less well tolerated than placebo; for discontinuation, the combination of fluoxetine and olanzapine, valproate, olanzapine alone and lamotrigine ranked higher than placebo. When response for all randomised participants (that is, assuming the dropouts did not respond) were compared, moclobemide and ziprasidone were also ranked below placebo. Other interventions that were included in the network but were not statistically superior to placebo were imipramine, lithium, moclobemide, paroxetine and ziprasidone. Excluding valproate, which only 48 people received, the five efficacious interventions were received by 292 to 1867 participants.

Table 9: Pharmacological interventions for acute episodes of bipolar depression (results from network meta-analysis)

Intervention	N	Response 1	Conditional response2	Discontinuation	Study ID(s)
Aripiprazole	385	0.41 (0.04, 3.38)	0.17 (0.00, 5.97)	1.58 (1.09, 2.31)	BRISTOLMYERSSQUIB2006, BRISTOLMYERSSQUIB2007, QUANTE2010
Fluoxetine and olanzapine	292	2.25 (1.58, 3.18)	2.37 (1.37, 4.29)	0.66 (0.43, 0.99)	BROWN2006, TOHEN2003,
Imipramine	111	1.06 (0.43, 2.48)	1.67 (0.49, 6.02)	1.36 (0.56, 3.37)	NEMEROFF2001, SILVERSTONE2001,
Lamotrigine	810	1.42 (1.13, 1.77)	1.44 (1.07, 2.00)	0.96 (0.74, 1.27)	BROWN2006, CALABRESE1999, CALABRESE2008d, CALABRESE2008c, CALABRESE2008b, CALABRESE2008a, VANDERLOOS2009,
Lithium	136	1.35 (0.88, 2.07)	1.77 (0.95, 3.32)	1.03 (0.60, 1.74)	YOUNG2010
Lurasidone	518	2.15 (1.58, 2.94)	3.00 (1.92, 4.72)	1.16 (0.78, 1.74)	SUNOVION2012a, SUNOVION2012b

Moclobemide	81	0.78 (0.26, 2.20)	1.17 (0.25, 5.81)	1.66 (0.51, 5.46)	SILVERSTONE2001
Olanzapine	713	1.41 (1.09, 1.83)	1.54 (0.98, 2.45)	0.86 (0.61, 1.20)	TOHEN2003, TOHEN2012
Paroxetine	155	1.21 (0.81, 1.80)	1.38 (0.77, 2.51)	0.97 (0.60, 1.51)	MCELROY2010, NEMEROFF2001,
Quetiapine	1867	1.69 (1.39, 2.06)	2.59 (1.94, 3.55)	1.03 (0.82, 1.29)	CALABRESE2005, MCELROY2010, SUPPES2010, THASE2006, YOUNG2010,
Valproate	48	2.7 (1.08, 7.56)	3.37 (1.07, 11.02)	0.62 (0.26, 1.45)	DAVIS2005, GHAEMI2007, MUZINA2011,
Ziprasidone	675	0.99 (0.77, 1.26)	1.27 (0.87, 1.91)	1.44 (1.06, 1.96)	PFIZER2009a, PFIZER2009b, SACHS2011,

Note. All effects (median OR and 95% CI) compared with placebo (N = 3215), which was included in BRISTOLMYERSSQUIB2006, BRISTOLMYERSSQUIB2007, CALABRESE1999, CALABRESE2005, CALABRESE2008a, CALABRESE2008b, CALABRESE2008c, CALABRESE2008d, DAVIS2005, GHAEMI2007, MCELROY2010, MUZINA2011, NEMEROFF2001, PFIZER2009a, PFIZER2009b, QUANTE2010, SACHS2011, SUNOVION2012a, SUNOVION2012b, SUPPES2010, THASE2006, TOHEN2003, TOHEN2012, VANDERLOOS2009, YOUNG2010.

1Effect calculated using the number of participants randomised to treatment as the denominator.
2Effect calculated using the number or participants who did not discontinue treatment as the denominator.

7.3.2.4 Pharmacological interventions for acute episodes of bipolar depression that could not be included in the network meta-analysis

One RCT (N = 174; POST2006) published in 2006 compared bupropion, sertraline and venlafaxine in outpatients. In the total sample, mean age was 42 years, 50% were female and 73% were diagnosed with bipolar I disorder. Little difference was found between any of the groups on response and discontinuation.

7.3.2.5 Nutritional interventions for acute episodes of bipolar depression

One RCT (N = 116) published in 2006 compared medication as usual with or without eicosapentaenoic acid supplementation (KECK2006b (Keck et al., 2006). There was very low quality evidence that eicosapentaenoic acid supplementation was not associated with a reduction in depressive symptoms (see Appendix X).