Appendix Hoofdstuk 7c: Kosteneffectiviteit van interventies voor manie en bipolaire depressie

In deze appendix staan de uitkomsten beschreven van onderzoek naar kosteneffectiviteit. Het deel over bipolaire depressie is bewerkt voor de Nederlandse situatie.

Acute mania: Health economics evidence

Systematic literature review

The systematic search of the economic literature undertaken for the guideline identified no study on the cost effectiveness of nutritional interventions and 4 eligible studies on the cost effectiveness of pharmacological treatments for adults with bipolar disorder in a manic, hypomanic or mixed episode (Bridle et al., 2004; Caro et al., 2006; Revicki et al., 2003; Zhu et al., 2005). Of these, only the study by Bridle and colleagues was conducted in the UK, while the rest three studies were conducted in the US.

Olanzapine versus valproate semisodium

Revicki and colleagues (2003) evaluated the cost effectiveness of valproate semisodium versus olanzapine in adults with bipolar I disorder in a manic episode in the US. The economic analysis was conducted alongside a multi-centre RCT (ZAJECKA2002). The study was a cost consequence analysis; the RCT outcomes considered in the analysis were the participants' clinical improvement based on the Mania Rating Scale (MRS) from the Schedule for Affective Disorders and Schizophrenia (SADS) Change Version and the Hamilton Rating Scale for Depression (HAM-D), and the participants' Health Related Quality of Life (HRQoL) measured by the Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q) and the number of days with restricted activity. The perspective of the analysis was that of a third-party payer. Costs included hospitalisation costs, physicians' fees, costs of emergency room, costs of psychiatric, physician, psychologist or other mental health provider visits, home health service visit costs and medication costs. HRQoL and resource use data were collected via telephone interviews; a number of resource use data, such as the number of inpatient physician visits and type of outpatient visits, were based on assumptions. National unit costs were used. The time horizon of the analysis was 12 weeks. Participants in the RCT discontinued treatment if they did not improve after 3 weeks, but data were still collected for a total period of 12 weeks.

The results of the analysis showed that there were no significant differences between the two drugs in terms of clinical, HRQoL and economic outcomes over the 12-week period. Valproate semisodium was associated with significantly lower outpatient costs compared with olanzapine; nevertheless, total direct medical costs associated with the two drugs were similar (mean total cost per

person $13,703 for valproate semisodiumand$15,180 for olanzapine, p = 0.88, cost year not stated). The study is partially applicable to the UK context as it was conducted in the US. Moreover, it is characterised by potentially serious limitations, relating to the short time horizon of the analysis (12 weeks), the use of assumptions for some resource use data, and potential conflicts of interest.

Zhu and colleagues (2005) also conducted a cost consequence analysis alongside a multi-centre RCT (TOHEN2002) to evaluate the cost effectiveness of olanzapine versus valproate semisodium in adults with bipolar I disorder that were hospitalised for a manic or mixed episode in the US. The time horizon of this analysis was 47 weeks, comprising 3 weeks of acute phase and 44 weeks of maintenance phase. Only participants who entered the maintenance phase of the RCT were included in the economic analysis (59% of the initial study sample). The clinical outcomes considered were the clinical improvement based on the Young Mania Rating Scale (YMRS) and the rate of symptom remission (defined as YMRS score ≤12) at 3 weeks, and the median time to remission of manic symptoms. The perspective of the analysis was that of a third-party payer. Cost elements included hospitalisation (full and partial), outpatient psychiatric physician and other mental health provider visits, emergency room visits, home visits by healthcare professionals, medication and laboratory tests. Effectiveness and resource use data were taken from the RCT; resource use data were collected from hospital and other medical records and family reports. National unit costs were used.

According to the analysis, total costs were similar between the two drugs (mean total cost per person $14,967 for olanzapine, $15,801 for valproate semisodium, p > 0.05, cost year 2000). Olanzapine was found to be significantly better than valproate semisodium in improving manic symptoms at 3 weeks and in the percentage of people achieving remission (54.4% versus 42.3%, respectively). The median time to remission was 14 days for olanzapine and 62 days for valproate semisodium. The results of the analysis suggest that olanzapine is a more effective treatment option that valproate semisodium for people with bipolar disorder experiencing mania at no extra cost. The study is partially applicable to the NHS context as it was conducted in the US. Moreover, it is characterised by potentially serious limitations including the design of the study regarding collection of resource use data and potential conflicts of interest.

Quetiapine versus usual care

Caro and colleagues (2006) developed a discrete event simulation model to evaluate the cost effectiveness of quetiapine versus usual care in adults with bipolar I disorder experiencing a manic episode in the US. Usual care comprised 45% monotherapy with lithium, 25% lithium plus risperidone, 25% lithium plus olanzapine, and 5% lithium plus quetiapine. The time horizon of the analysis was 100 days. The analysis adopted a third-party payer perspective. Cost elements consisted of hospitalisation and physician fees, emergency room and intensive

care units, routine physician and psychiatrist visits, laboratory tests, medication and management of side effects. The outcome measures used were the percentage of people responding at 21 days and the percentage of people remitting at 84 days. Clinical data for the economic model were taken from a literature review, whereas resource use data were derived from administrative databases; national unit costs were used.

Quetiapine was found to be overall less costly than usual care (mean total cost per person $5,525 for quetiapine and $6,912 for quetiapine in 2004 prices). It was also found to be more effective than usual care: the percentage of people responding at 21 days was 54% for quetiapine and 43% for usual care; the percentage of people remitting at 84 days was 80% for quetiapine and 74% for usual care. Consequently quetiapine was the dominant treatment option. Results were sensitive to drug prices, discharge criteria and side-effect management costs. The study is partially applicable to the UK context as it was conducted in the US; the definition of usual care may not reflect usual care in the UK. The analysis is characterised by a number of potentially serious limitations including the source of cost and effectiveness data and potential conflicts of interest.

Antipsychotic drugs (olanzapine, quetiapine and haloperidol) compared with lithium and valproate semisodium

The economic analysis by Bridle and colleagues (2004) was the only study undertaken in the UK. The objective of the study, which informed a previous NICE Technology Appraisal on the use of newer anti-manic drugs (NICE, 2003), was to evaluate the cost effectiveness of quetiapine, olanzapine and valproate semisodium in the treatment adults with bipolar disorder experiencing an manic episode. The study was based on decision-analytic modelling. Effectiveness data were derived from a systematic review and network meta-analysis. The availability of effectiveness data in the network meta-analysis determined the choice of drugs included in the economic analysis. The following drugs were thus considered in the analysis: quetiapine, olanzapine, valproate semisodium, haloperidol and lithium.

The primary measure of outcome was the number of responders to treatment; response was defined as ≥50% improvement in manic symptoms, expressed in changes in YMRS scores. The time horizon was equal to 3 weeks in the base-case analysis, to reflect the most commonly reported length of follow-up for which effectiveness data were provided in the clinical trials. Estimated costs, expressed in 2001–2002 prices, included direct medical costs from the NHS perspective; these consisted of hospitalisation and drug-acquisition costs, as well as costs of diagnostic and laboratory tests required for monitoring. Resource use data were based on expert opinion, information from manufacturers and further assumptions. Unit costs were taken from national sources. Costs of treating adverse events were not included in the analysis, because of lack of relevant data reported in the literature. However, the authors' opinion was that the

majority of adverse events associated with the drugs compared were unlikely to have significant resource use implications in the 3-week time horizon of the model. Hospitalisation costs were estimated to be the same for all drug treatment options, as all people experiencing a manic episode were assumed to be hospitalised at the start of the model and to remain hospitalised for the total 3-week period, regardless of response to treatment.

The base-case results of the analysis showed that mean response rates for olanzapine (0.54) and haloperidol (0.52) were higher than for lithium (0.50), quetiapine (0.47) and valproate semisodium (0.45). Haloperidol had the lowest mean total costs per person (£3,047) in comparison to valproate semisodium (£3,139), olanzapine (£3,161), lithium (£3,162) and quetiapine (£3,165). In terms of cost effectiveness, lithium, valproate semisodium and quetiapine were dominated by haloperidol as they were all less effective and more costly than haloperidol. Compared with haloperidol, olanzapine was more effective and resulted in higher total costs, demonstrating an incremental cost effectiveness ratio (ICER) equal to£7,179 per additional responder. This means that if decision-makers are prepared to pay less than £7,179 per additional responder, then haloperidol is the optimal decision; however, if they are prepared to pay at least £7,179 per additional responder, then olanzapine is the most cost-effective option.

One-way sensitivity analyses showed that results relating to dominance of haloperidol were robust to alternative assumptions tested, such as discharge of non-responders at a later time than responders, treatment of non-responders with second and third-line pharmacological therapies, reductions in diagnostic and laboratory costs, inclusion of effectiveness data for people initially excluded from analysis according to a modified intention-to-treat approach, and inclusion of treatment costs for extrapyramidal symptoms because of haloperidol use. Under these scenarios, the ICER of olanzapine compared with haloperidol ranged between £1,236 (when longer hospitalisation was assumed for non-responders) and £7,165 (when second and third-line treatment was assumed for non-responders) per additional responder. Base-case results were sensitive only to the entire exclusion of diagnostic and laboratory costs from the analysis, which constituted a rather extreme scenario.

Probabilistic analysis demonstrated that, for a willingness to pay (WTP) equal to £20,000 per additional responder, the probabilities of each drug being cost-effective were: olanzapine 0.44, haloperidol 0.37, lithium 0.16, quetiapine 0.02 and valproate semisodium 0.01. The probability that olanzapine was cost-effective increased as the WTP increased: for a maximum WTP £10,000 per additional responder this probability reached 0.42, increasing to 0.45 if the maximum WTP rose to £40,000. At the extreme of a zero value placed on the WTP for an additional responder, haloperidol was the most cost-effective option (with probability equalling 1).

Although the study was conducted in the UK, it is only partially applicable to the NICE context because its primary measure of outcome was the rates of response and not the quality-adjusted life year (QALY), which is the preferred outcome measure by NICE, due to lack of appropriate utility data. As a result, the reported ICERs are difficult to interpret as there is no set threshold for the WTP per additional responder to anti-manic therapy. In addition, although the study was well conducted, it is characterised by potentially serious limitations: first of all, the model had a very short time horizon of 3 weeks, which was nevertheless dictated by the time horizon of the RCTs included in the network meta-analysis. This means that potential differences across drugs regarding benefits and resource use, including the overall length of hospitalisation (beyond 3 weeks), were not taken into account. However, potential differences in the length of hospitalisation among drugs may affect significantly their relative cost effectiveness, as inpatient care is the major driver of total medical costs associated with treatment of mania. Cost differences between drugs were found to be very small and were attributed exclusively to differences in acquisition and monitoring costs, as hospitalisation costs were assumed to be the same across drugs over the time period of 3 weeks. Finally, omission of costs and HRQoL aspects of side effects from the analysis was also acknowledged by the authors as a further limitation of their study.

Overall conclusions from existing economic evidence

The existing economic evidence on drugs for the treatment of mania in people with bipolar disorder is rather limited and not directly applicable to the NICE decision-making context. All studies included in the review are characterised by potentially serious limitations. Evidence from the US suggests that olanzapine and valproate semisodium are associated with similar overall costs; in terms of effectiveness one study showed superiority of olanzapine, and the other study found no difference in effectiveness. Another US study indicated that quetiapine was dominant (more effective and less costly) than usual care. The only UK study included in the review showed that haloperidol was dominant over lithium, valproate semisodium and quetiapine. Olanzapine was more effective and more costly than haloperidol, with an ICER equal to £7,179 per additional responder. However, the study is characterised by potentially serious limitations and its results are not easy to interpret due to lack of use of QALYs as a measure of outcome.

It needs to be noted that quetiapine and olanzapine are now available in generic form, and therefore their acquisition cost is lower than the cost of the patented forms evaluated in the studies included in the systematic review. Thus their relative cost effectiveness is likely higher than that suggested in the literature.

Economic modelling

Introduction – objective of economic modelling

The cost effectiveness of pharmacological interventions for the treatment of adults with bipolar disorder experiencing a manic episode was identified by the GDG as an area with potentially major resource use implications that should be addressed by economic modelling. However, the availability of clinical and cost data did not allow the development of a model with a time horizon longer than 3 weeks that would overcome the limitations characterising the study by Bridle and colleagues (2004). Therefore, a simple economic analysis was attempted, which updated the costs and clinical data reported by Bridle and colleagues (2004) and allowed the GDG to consider the costs associated with pharmacological interventions for mania alongside their clinical effectiveness as reported in Cipriani and colleagues (2011). In addition, a cost-utility analysis was conducted, using available utility data that allowed outcomes to be expressed in the form of QALYs.

Economic modelling methods

Interventions assessed

The interventions that were assessed in this economic analysis were determined by the availability of data reported in the network meta-analysis by Cipriani and colleagues (2011). Only drugs that were found to be effective in this study and licensed in the UK were considered in the economic analysis. Cipriani and colleagues (2011) evaluated the following drugs: aripiprazole, asenapine, carbamazepine, valproate, gabapentin, haloperidol, lamotrigine, lithium, olanzapine, quetiapine, risperidone, topiramate and ziprasidone. Paliperidone was not assessed separately, but relevant data were pooled with risperidone data, as paliperidone is the main active metabolite of risperidone. The economic analysis did not consider ziprasidone, because this is not licensed in the UK. Moreover, gabapentin, lamotrigine and topiramate were found to be not significantly better than placebo in the network meta-analysis and were thus excluded from the economic analysis. Thus the economic analysis assessed the costs and outcomes of the following nine drugs: aripiprazole, asenapine, carbamazepine, valproate, haloperidol, lithium, olanzapine, quetiapine and risperidone.

Costs and outcomes considered in the analysis

The economic analysis adopted the NHS and personal social services (PSS) perspective, as recommended by NICE (2012). Costs included hospitalisation costs, drug acquisition costs and costs of laboratory testing. The measures of effectiveness were determined by the outcome measures reported in Cipriani and colleagues (2011), which included the change scores on the YMRS as a primary outcome, and the proportion of people who responded to treatment as a secondary outcome. Moreover, the economic analysis estimated the number of QALYs gained associated with each pharmacological treatment.

Time horizon of the analysis

The time horizon of the economic analysis was 3 weeks, the same as in the study by Bridle and colleagues (2004), which reflected the time horizons of the RCTs included in the network meta-analysis that provided the effectiveness data.

Clinical input parameters

All clinical input parameters were taken from the study by Cipriani and colleagues (2011). These included the SMDs of YMRS scores and the ORs of response rates, as well as the baseline probability of response for placebo. The latter was estimated by pooling the data from all placebo arms included in the network meta-analysis and found to equal 31.1%. This baseline probability of response was used in order to estimate the probability of response for each drug using the following formulae:

$$px = oddsx/ (1 + oddsx)$$

and

$$oddsx = (1/ORb,x)* pb/(1-pb)$$

where pb the probability of response for placebo (baseline), ORb,x the odds ratio for response of placebo versus each drug as reported in Cipriani and colleagues (2011) and oddsx the odds of each drug to achieve response.

Utility data and estimation of quality-adjusted life years

In order to express outcomes in the form of QALYs, the health states of the economic model need to be linked to appropriate utility scores. Utility scores represent the HRQoL associated with specific health states on a scale from 0 (death) to 1 (perfect health). More details on the estimation of utility scores, the NICE criteria on selection of available utility data and on the systematic review of the literature that aimed to identify utility scores associated with distinct health states experienced by adults with bipolar disorder are provided in section 1.4.5. This analysis considered utility scores corresponding to the health states of 'mania' equalling 0.44, and 'full response – euthymia' equalling 0.90, as reported in Table 12; the difference in utility between these states (0.46) was estimated using data reported in Revicki and colleagues (2005). The utility score for mania was used for all people at the start of the model and for people not responding to treatment; the utility score for euthymia was used for people responding to treatment. The model assumed linear increase in utility in those responding to treatment between the start of the model and the point where response was achieved.

Cost data

Similar to the economic analysis by Bridle and colleagues (2004), people in all arms of the economic model were assumed to be hospitalised over the 3-week time horizon of the analysis. Therefore, hospitalisation costs were the same across all drugs and were excluded from the guideline analysis.

The drug daily dosage was determined according to optimal levels of administration (based on the BNF and the GDG expert opinion) and was consistent with the dosage range reported in the RCTs included in the network meta-analysis by Cipriani and colleagues (2011). Drug acquisition costs were taken from the NHS Electronic Drug Tariff, February 2014 (NHS Business Services Authority, 2014).

Required laboratory testing was determined by the GDG expert opinion. It was agreed that at initiation of all drugs a number of tests should be undertaken, including electrocardiogram (ECG), assessment of renal function (creatinine, blood urea and electrolytes), glucose, lipid profile and thyroid function tests. The costs of these tests were not included in the analysis, as they were common to all arms of the model. In addition to these tests, the GDG expressed the opinion that liver function should be tested at initiation of all drugs except lithium; for lithium, 3 tests of serum lithium concentration were required to determine optimal dose. The cost of liver function testing was taken from data reported in the economic analysis described in the previous NICE guideline (NCCMH, 2006). The cost of serum lithium concentration testing was taken from the Newcastle upon Tyne Hospitals NHS trust biochemistry laboratory services tariff for 2006-7.

All costs were uplifted to 2014 prices using the Hospital and Community Health Services (HCHS) pay and prices inflation index (Curtis, 2013). The inflation index for the year 2014 was estimated using the average value of the HCHS pay and prices indices of the previous 3 years.

The drug daily dosages and the associated acquisition costs, as well the laboratory testing costs that were utilised in the model are reported in Table 1.

Table 1. Average daily dosage, daily and 3-week acquisition costs, and additional required laboratory testing costs of pharmacological interventions for the treatment of adults with bipolar disorder experiencing a manic episode included in the economic analysis (2014 prices)

Drug	Daily dosage	Daily drug cost	3-week drug cost	Laboratory test and cost
Aripiprazole	15 mg	£6.86	£144.06	Liver function: £4.37
Asenapine	10 mg twice daily	£3.42	£71.82	Liver function: £4.37
Carbamazepine	500 mg	£0.32	£6.77	Liver function: £4.37
Valproate	1500 mg	£0.97	£20.41	Liver function: £4.37
Haloperidol	5 mg twice daily	£0.23	£4.76	Liver function: £4.37

Lithium	1400 mg	£0.12	£2.59	Lithium concentration: 3 x £3.25
Olanzapine	15 mg	£0.08	£1.61	Liver function: £4.37
Quetiapine	300 mg twice daily	£0.17	£3.55	Liver function: £4.37
Risperidone	4 mg	£0.04	£0.79	Liver function: £4.37
Drug acquisition costs from the NHS Electronic Drug Tariff, February 2014 (NHS Business Services Authority, 2014). Liver function testing cost from (NCCMH, 2006). Serum lithium concentration testing cost from the Newcastle upon Tyne Hospitals NHS trust biochemistry laboratory services tariff for 2006-7.				

Data analysis

Estimated costs of pharmacological interventions are presented alongside effectiveness data (SMDs of YMRS scores and ORs of response as reported in Cipriani and colleagues (2011)) and the mean QALY gain per person. Formal synthesis of costs and SMDs in an ICER was not attempted, as the resulting figures would be difficult to interpret and therefore would not be useful in decision-making. On the other hand, ICERs expressing cost per additional responder were estimated despite the fact that they were difficult to interpret, to enable comparisons with the results reported in Bridle and colleagues (2004). In addition, incremental analysis where the ICER was expressed as cost/QALY was undertaken. Probabilistic analysis was not possible to undertake using the summarised efficacy data (mean and 95% CIs) that were reported in Cipriani and colleagues (2011). The cost data used in this analysis were very limited and were not subject to uncertainty, as the drug and laboratory testing unit prices are determined. Therefore, other sensitivity analysis was not attempted.

Economic modelling results

Results of the economic analysis using the SMDs and the ORs of response of each drug versus placebo are presented in Table 2 and Table 3, respectively. Table 3 also presents the QALY gains per person associated with each drug. In both tables, drugs have been ordered from the most to the least effective. As shown in Table 2, the 3 most effective drugs in terms of SMD are haloperidol, risperidone and olanzapine; these drugs have also the lowest costs, all below £10 per person. These drugs are followed by quetiapine and lithium, which have comparable costs, as well as aripiprazole, which, however, has a total acquisition and laboratory testing cost of £148.

Table 2. Results of the economic analysis of pharmacological interventions for the treatment of adults with bipolar disorder experiencing a manic episode: effectiveness expressed by the

standardised mean difference (SMD) of YMRS scores compared with placebo and costs

Drug	Effectiveness: SMD Mean (95% CIs)	Cost per person
Haloperidol	–0.56 (–0.68 to –0.43)	£9.12
Risperidone	–0.50 (–0.63 to –0.38)	£5.16
Olanzapine	–0.43 (–0.54 to –0.32)	£5.97
Quetiapine	–0.37 (–0.51 to –0.23)	£7.92
Lithium	–0.37 (–0.50 to –0.25)	£12.34
Aripiprazole	–0.37 (–0.51 to –0.23)	£148.43
Carbamazepine	–0.36 (–0.60 to –0.11)	£11.14
Asenapine	–0.30 (–0.53 to –0.07)	£76.19
Valproate	–0.20 (–0.37 to –0.04)	£24.77

In terms of ORs of response and QALYs, the 4 most effective drugs were carbamazepine, haloperidol, olanzapine and risperidone, all with comparable costs. These are followed by quetiapine, which has also comparable costs, valproate, which has somewhat higher costs, and aripiprazole, which is by far the most costly drug of the analysis. According to formal incremental analysis, all drugs below the 4 most effective drugs are dominated by absolute dominance, as they are less effective and more costly than one of more of the 4 most effective drugs. Haloperidol and olanzapine are dominated by rules of extended dominance (the latter occurs when an option is less effective and more costly than a linear combination of two alternative options). The ICER of carbamazepine versus risperidone is £149 per additional responder or £11,191/QALY. It needs to be noted that carbamazepine was not among the most effective drugs in the analysis of YMRS change scores, which was the primary analysis of efficacy data in Cipriani and colleagues (2011). If carbamazepine is excluded from incremental analysis, then haloperidol and olanzapine are not dominated anymore. The ICER of haloperidol versus olanzapine is £283 per additional responder or £21,363/QALY and the ICER of olanzapine versus risperidone is £151 per additional responder or £11,412/QALY. Using the NICE cost effectiveness threshold of £20,000-£30,000/QALY, olanzapine becomes the most cost-effective option if carbamazepine is excluded from analysis. This is followed by haloperidol (ICER versus risperidone £240 per additional responder or £18,119/QALY) and risperidone. Quetiapine is the next most cost-effective option, as it dominates all the remaining drugs in the analysis.

The ICERs expressing cost per additional responder are difficult to interpret, as there is no set threshold regarding the WTP per additional responder to treatment for mania. Nevertheless, they were estimated to enable comparison with respective ICERs reported in Bridle and colleagues (2004). The comparison reveals that the ICERs estimated in this analysis are much lower than those reported by Bridle and colleagues, who estimated an ICER of olanzapine versus haloperidol equal to £7,179 per additional responder; this discrepancy may be attributable to the very different drug acquisition costs between the guideline

analysis and the analysis by Bridle and colleagues (2004), as, since the latter, many of the drugs considered have become available in generic form. It should also be noted that the total costs reported in this analysis are substantially lower than those reported by Bridle and colleagues (2004), because this analysis did not include costs of hospitalisation, which were common across all arms and were thus cancelled out.

Table 3. Results of the economic analysis of pharmacological interventions for the treatment of adults with bipolar disorder experiencing a manic episode: effectiveness expressed by the odds ratios (ORs) of response rates of placebo versus each drug, QALYs, costs and incremental cost effectiveness ratios

Drug	Effectiveness: OR Mean (95% CIs)	Probability of response	QALYs / person	Cost / person	ICER
Carbamazepine	0.40 (0.22 to 0.77)	0.530	0.0324	£11.14	£149/extra responder £11,191/QALY
Haloperidol	0.44 (0.33 to 0.58)	0.506	0.0321	£9.12	£283/extra responder £21,363/QALY - dominated by ED
Olanzapine	0.46 (0.36 to 0.58)	0.495	0.0320	£5.97	£151/extra responder £11,412/QALY - dominated by ED
Risperidone	0.47 (0.35 to 0.61)	0.490	0.0319	£5.16	
Quetiapine	0.50 (0.37 to 0.66)	0.474	0.0317	£7.92	Dominated
Valproate	0.50 (0.36 to 0.70)	0.474	0.0317	£24.77	Dominated
Aripiprazole	0.50 (0.38 to 0.66)	0.474	0.0317	£148.43	Dominated
Lithium	0.55 (0.38 to 0.79)	0.451	0.0314	£12.34	Dominated
Asenapine	0.59 (0.31 to 1.13)	0.433	0.0311	£76.19	Dominated

ED = extended dominance

The methodology checklist and the economic evidence profile of the analysis are provided in Appendix 30 and Appendix 32, respectively.

Discussion – limitations of the analysis

The results of the economic analysis suggest that haloperidol, olanzapine, risperidone and quetiapine may be more cost-effective options compared with the other drugs assessed in the analysis. Carbamazepine was shown to be the most effective (and cost-effective) option when ORs of response and QALYs were used, but not in the analysis that utilised SMDs. After excluding carbamazepine from the cost-utility analysis, olanzapine became the most cost-effective

treatment option, followed by haloperidol, risperidone and quetiapine. It has to be noted that the efficacy and cost differences between haloperidol, olanzapine, risperidone and quetiapine were overall shown to be rather unimportant.

The economic analysis is very simplistic and has taken into account only costs associated with drug acquisition and additional laboratory tests required for each drug over a period of 3 weeks. This short time horizon was imposed by the short time horizons of the RCTs that were included in the network meta-analysis that provided the effectiveness data. Side effects and their impact on costs and HRQoL were not considered in the analysis, due to the short time horizon and the lack of relevant data. Hospitalisation costs were assumed to be the same for all drugs over 3 weeks, as all people with bipolar disorder experiencing an acute episode were estimated to be hospitalised over the first 3 weeks of acute treatment. However, the total length of hospitalisation and outcomes of drugs beyond 3 weeks were not taken into account in the analysis due to lack of relevant data. If some drugs result in better outcomes beyond the period of the 3 weeks and reduce the total length of hospitalisation, then they are expected to be more cost-effective, as hospitalisation is the most substantial driver of costs in the treatment of mania (the mean cost of Mental Health Care Clusters per bed-day was £344 in 2013, according to NHS reference costs (NHS, 2013)).

Another limitation of the analysis is the use of utility data from Revicki and colleagues (2005) owing to the lack of more relevant utility data for the state of mania. The study described hypothetical health states using vignettes, which were valued by stable outpatients with bipolar disorder in the US. As discussed in section 1.3.7, these utility values do not meet NICE criteria on use of utility values and do not reflect the UK general population's preferences. The results of the cost-utility analysis should be therefore interpreted with caution.

Overall conclusions from economic evidence

The existing economic evidence is rather limited and not directly applicable to the NICE decision-making context; all studies are characterised by potentially serious limitations. In the economic analysis conducted for this guideline, haloperidol, olanzapine, risperidone and quetiapine appear to be more cost-effective options than other drugs included in the analysis. However, the analysis has not overcome many of the limitations characterising previous studies. Factors such as acceptability, rate and type of side effects associated with each drug should be considered when making recommendations.

Bipolar Depression: Health economics evidence

Systematic literature review

The systematic search of the economic literature undertaken for the guideline identified one eligible study on the cost effectiveness of pharmacological interventions (Ekman et al., 2012) and one eligible study on the cost

effectiveness of nutritional interventions (Cheema et al., 2013) for adults with bipolar disorder in an acute depressive episode.

The study by Ekman and colleagues (2012) assessed the cost effectiveness of quetiapine versus a number of pharmacological treatment options in adults with bipolar disorder (I or II) in the UK. The study was based on decision-analytic modelling. Two separate analyses were undertaken: one where the study population entered the model in an acute episode of bipolar depression, and another one where the study population entered the model in remission. Both analyses had a 5-year time horizon and considered the following treatment options: quetiapine; quetiapine added to a mood stabiliser (lithium or valproate semisodium); olanzapine; olanzapine plus lithium, with olanzapine replaced by venlafaxine in acute depression; olanzapine plus lithium, with olanzapine replaced by paroxetine in acute depression; aripiprazole that was replaced by olanzapine and venlafaxine in acute depression; and a mixed scenario where risperidone was administered in mania, venlafaxine and lithium were administered in acute depression, and olanzapine was administered as maintenance treatment.

The study adopted the NHS perspective. Costs included hospitalisation costs, costs of outpatient care, costs associated with crisis teams, staff costs (senior house officer, GP, community psychiatric nurse, practice nurse, dietician), drug acquisition costs, laboratory test costs, and costs of adverse events. Indirect costs (productivity losses) were considered in a sensitivity analysis. The measure of outcome was the QALY. Relative effects across drugs were taken from RCTs and published meta-analyses of trials. Resource use data were taken from published sources, which, however, reported estimates based on expert opinion. Unit costs were taken from national sources.

The study is directly applicable to the UK. However, evidence synthesis was based on indirect comparisons between drugs, using placebo as baseline; however, as the authors acknowledged, the meta-analyses used to derive the relative effects were not similar in terms of the phase of the disorder examined and the measures of outcome used. Moreover, it is not clear whether the study populations and designs across all RCTs used in evidence synthesis (including those considered in the published meta-analyses) were similar enough to allow indirect comparisons of drugs. Overall, it appears that methods of evidence synthesis were inappropriate, introducing bias in the economic analysis. For this reason, the study was judged to suffer from very serious limitations and was therefore not considered further when making recommendations.

Cheema and colleagues (2013) evaluated the cost effectiveness of ethyl-eicosapentaenoic acid (ethyl-EPA) adjunctive to mood stabilisers versus mood stabilisers alone in adults with bipolar I disorder in a stable (euthymic) state, from the perspective of the UK NHS. The study, which was based on decision-analytic modelling, is described here because it has utilised effectiveness data

from a 12-week RCT that assessed the efficacy of ethyl-EPA in people with bipolar depression (FRANGOU2006). This RCT was excluded from the guideline systematic review because participants were not acutely depressed. The economic analysis extrapolated the efficacy data from this trial to stable adults with bipolar disorder experiencing acute episodes, over 1 year; efficacy of ethyl-EPA in reducing depressive symptoms over 12 weeks was assumed to correspond to efficacy in preventing acute manic and depressive episodes over 1 year. This was considered a very serious limitation of the analysis; consequently the study was not considered further when formulating guideline recommendations.

Economic modelling

Introduction – objective of economic modelling

The cost effectiveness of pharmacological interventions for adults with bipolar disorder experiencing an acute depressive episode was considered by the GDG as an area with likely significant resource implications. Existing economic evidence in this area was limited to one study that was conducted in the UK. The study was characterised by potentially serious limitations and did not assess the whole range of interventions that are available in the UK for the treatment of acute depression in adults with bipolar disorder. The clinical evidence in this area was judged to be sufficient and of adequate quality to inform primary economic modelling. Based on the above considerations, this area was prioritised for further economic analysis. An economic model was therefore developed to assess the relative cost effectiveness of pharmacological interventions for adults with bipolar disorder experiencing an acute depressive episode in the UK, which was then adapted to the Netherlands by adjusting prices and resource use where necessary.

Economic modelling methods

Interventions assessed

The guideline economic analysis assessed pharmacological interventions that were included in the relevant network meta-analysis conducted for this guideline. The economic model considered interventions that were found to be effective in the network meta-analysis and are available in the UK. Aripiprazole was excluded from the economic analysis, since the network meta-analysis indicated that it is ineffective in the treatment of acute depression in adults with bipolar disorder. Lurasidone and ziprasidone were not considered in the economic analysis because they are not available both in the UK and in the Netherlands.

Based on the above criteria the following pharmacological interventions were included in the economic analysis: imipramine, lamotrigine, lithium,

moclobemide, olanzapine, paroxetine, quetiapine, valproate semisodium, and the combination of fluoxetine and olanzapine.

The model also considered no pharmacological treatment (reflected in treatment with placebo) consisting, in terms of resource use, of visits to healthcare professionals only, in order to assess the cost effectiveness of active interventions versus a non-specific medical management (used as a benchmark).

Model structure

A decision-analytic model in the form of a decision-tree was constructed using Microsoft Office Excel 2010. The model estimated the total costs and benefits associated with provision of each of the 10 treatment options (including no pharmacological treatment) to adults with bipolar disorder experiencing an acute depressive episode. The structure of the model, which aimed to simulate the course of acute bipolar depression and relevant clinical practice in the UK, was also driven by the availability of clinical data. The model was later adapted to represent the situation in the Netherlands.

According to the model structure, hypothetical cohorts of adults with bipolar disorder in acute depression were initiated on each of the 10 treatment options assessed. People initiated on a pharmacological treatment option could either continue treatment for 6 weeks or discontinue for any reason (for example because of intolerable side effects). Drug discontinuation was estimated to occur on average at 3 weeks from initiation of drug treatment. At the end of 6 weeks, people continuing treatment either responded to treatment fully or partially, or they did not respond. Assessment of response was undertaken at this point because 6 weeks was the median (and mode) time horizon of the studies considered in the guideline network meta-analysis that provided the response data for the model. People who responded to the initiated drug fully or partially continued their drug treatment for another 12 weeks at the same dosage, at the end of which they either experienced a manic or depressive relapse or did not relapse.

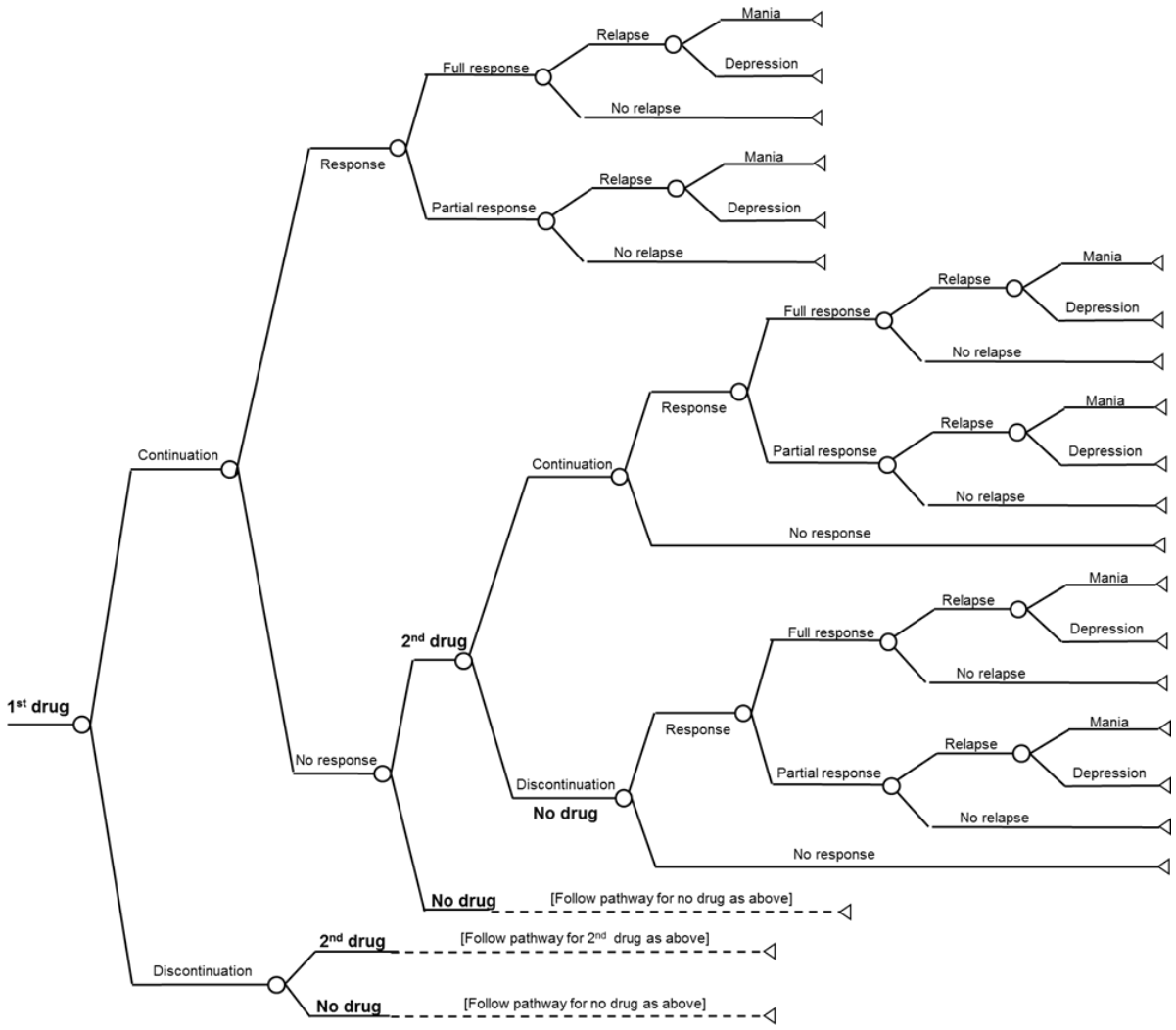
People discontinuing their initiated drug treatment at 3 weeks or not responding to this treatment after 6 weeks either stopped drug treatment (that is, they moved to no pharmacological treatment) or moved to a second drug treatment option; this was assumed to be a non-weighted ‘average’ mixture of all other drug treatment options assessed in the economic analysis (in terms of intervention costs and clinical outcomes), excluding the initiated drug treatment option. People initiated on the combination of fluoxetine and olanzapine could move to a mixture of all other drugs evaluated in the model except monotherapy with olanzapine, since the combination of the latter with fluoxetine had already failed. People under the second drug treatment option either continued the drug treatment or discontinued after 3 weeks and moved to no pharmacological treatment. Those continuing the second drug followed the same pathway as

people who continued the first drug (that is, no response or response, either full or partial, 6 weeks later, after which they could relapse to a manic or depressive episode or not relapse). People receiving a second drug treatment and not discontinuing remained on this drug for the remaining of the time horizon, whether they responded to this treatment or not.

People under no pharmacological treatment (either as initial treatment, or following discontinuation of, or no response to, their initiated drug treatment option) either responded to treatment, fully or partially, and could experience a manic or depressive relapse, or did not respond to treatment.

The time horizon of the analysis was 18 weeks, which consisted, for people responding to their initiated drug, of 6 weeks of treatment until assessment of the clinical outcome (6 weeks was the median time horizon of trials considered in the guideline network meta-analysis), and another 12 weeks of continuation of the drug, prior to initiation of long-term pharmacological maintenance treatment. The GDG expressed the opinion that people with acute bipolar depression that show responsiveness to a drug normally continue the drug as acute treatment, and at full dosage, for another 8 weeks and then they either take the drug as long-term maintenance treatment at the same dosage, or they receive the drug at gradually reduced dosages over a period of another 4 weeks, during which they start long-term maintenance treatment with another drug. For simplicity purposes as well as for consistency across model arms (as some drugs in the model are not suitable for long-term maintenance treatment), it was assumed that all people responding to a drug received its full dosage for the remaining of the model. The 18-week time horizon enabled capturing the full course of acute drug treatment for people who responded at 6 weeks (6 + 8 + 4 weeks), and was long enough to allow moving to second drug treatment and assessing response in cases where the 6-week initiated drug treatment failed; the model did not extend beyond 18 weeks because this would mean that some people in the model (those who responded at 6 weeks) would start maintenance treatment whereas others would be still receiving acute treatment for their depressive episode. Maintenance treatment was not considered in the model due to lack of appropriate and relevant data that were required to populate a longer-term economic model, as discussed in Chapter 7. A schematic diagram of the decision-tree is presented in Figure 1.

Figure 1. Schematic diagram of the economic model constructed for the evaluation of the relative cost effectiveness of pharmacological interventions for acute depression in adults with bipolar disorder.



Costs and outcomes considered in the analysis

The economic analysis adopteda health care perspective by considering direct medical costs. Direct non-medical costs, such as travel expenses, as well as indirect non-medical costs, such as productivity losses, were not considered in the analysis. Direct medical costs consisted of drug acquisition costs, laboratory testing costs, healthcare professional visit costs, as well as costs of hospitalisation and Intensive Home Treatment teams (IHTTs) for a proportion of people not responding to treatment. The measure of outcome was the QALY.

Clinical input parameters

Clinical model input parameters consisted of the probabilities of discontinuation and conditional response (in those not discontinuing) following first and second treatment; the probability of response in people under no pharmacological treatment; the probability of moving to no pharmacological treatment following discontinuation or no response to first pharmacological treatment; the probability of partial response in those responding; the probability of relapse in those responding fully or partially; and the probability of a manic episode in those relapsing.

The probabilities of discontinuation and response in those not discontinuing were taken from the network meta-analysis conducted for this guideline, the methods of which are reported in Appendix 11. For the economic analysis the first 50,000 iterations undertaken in WinBUGS were discarded and another 300,000 were run, thinned by 30, so as to obtain 10,000 iterations that populated the economic model. The results of the network meta-analysis that were used to populate the economic model are provided in Table 8. The table shows the mean probability of discontinuation and conditional response (that is, response in those not discontinuing) for each intervention considered in the economic analysis at the end of treatment (6 weeks).

For no pharmacological treatment (placebo), the data on probability of discontinuation and conditional response were combined in order to provide an overall probability of response in those under no pharmacological treatment (placebo), since the probability of discontinuation was not meaningful in an economic model that assumed that people were already under no pharmacological treatment. Thus, people discontinuing placebo were counted as non-responders.

Table 4. Results of network meta-analysis that were utilised in the economic model: probability of discontinuation and conditional response in adults with acute bipolar depression at end of treatment.

Intervention	Mean probability of discontinuation (95% credible intervals)	Mean probability of conditional response (95% credible intervals)
Imipramine	0.41 (0.17 to 0.69)	0.64 (0.26 to 0.92)
Lamotrigine	0.33 (0.16 to 0.53)	0.62 (0.33 to 0.85)
Lithium	0.35 (0.16 to 0.58)	0.66 (0.35 to 0.89)
Moclobemide	0.45 (0.16 to 0.77)	0.56 (0.16 to 0.91)
Olanzapine	0.31 (0.15 to 0.51)	0.63 (0.34 to 0.87)
Paroxetine	0.33 (0.15 to 0.55)	0.61 (0.30 to 0.86)
Quetiapine	0.35 (0.18 to 0.55)	0.74 (0.48 to 0.91)
Valproate	0.25 (0.08 to 0.50)	0.77 (0.43 to 0.95)
Fluoxetine and olanzapine	0.26 (0.11 to 0.45)	0.72 (0.43 to 0.91)

The probability of discontinuation remained the same for each drug when used as second drug option. The probability of conditional response for each drug, however, was assumed to be lower when the drug was used as second option. This reduction in probability of conditional response was assumed to be the same across all drugs and was estimated using data from a longitudinal study on

adults with unipolar major depression receiving one to four successive pharmacological treatment options (Rush et al., 2006), owing to the lack of relevant data on people with bipolar disorder. The reduction in response was also applied to no pharmacological treatment (placebo) for people moving to it after discontinuation of, or no response to, a pharmacological treatment option. It was estimated that the probability of response of each treatment option used as second choice was 0.59 of the probability of response for this option if used as first choice.

The probability of moving to no pharmacological treatment following discontinuation of, or no response to, first pharmacological treatment was based on the GDG expert opinion; the GDG estimated that 25% of people discontinuing their first drug and 10% of people not responding to their first drug moved to no pharmacological treatment.

The probability of partial response in those responding to treatment was assumed to be the same across all treatments and was estimated based on data reported in a pragmatic trial that compared a mood stabiliser plus adjunctive antidepressant therapy versus a mood stabiliser plus a matching placebo in adults with acute bipolar depression (bipolar depression I or II) (Sachs et al., 2007). According to data reported in this trial, out of 366 participants with acute depression, 165 achieved either transient remission or durable recovery (defined as euthymia for a minimum of 8 weeks) following treatment. The percentage of people achieving a transient remission was 43.6% (72/165), and this figure was used in the model to represent the probability of partial response in those responding to treatment.

The probability of relapse following full or partial response was estimated based on data reported in a prospective naturalistic study that followed 223 adults with bipolar disorder I or II for up to 20 years (Judd et al., 2008). The study reported the probability of relapse to a major acute episode following full and partial recovery from a previous acute episode (which could be manic or depressive), and these data were used to model the probability of relapse at the end of the 18 weeks for all people in the model that had responded to treatment, taking into account that the point at which response occurred differed across the various pathways in each cohort, so that the probability of relapse at the end of 18 weeks, which was assumed to be time-dependent, differed across the various pathways, too.

The probability of a manic episode in those relapsing was also estimated using data reported in Judd and colleagues (2008). The study reported that in 126 people with bipolar disorder who had recovered from an acute depressive or manic episode and experienced a relapse, 66 had a major depressive episode (52.4%), 26 had a manic episode (20.6%) and 34 had a mixed/cycling polarity episode (27.0%). For simplicity, the GDG advised that half of the mixed/cycling episodes should be considered manic and half should be considered depressive,

resulting in a ratio of manic to depressive acute relapses 34.1:65.9, and a probability of a manic episode in those relapsing of 0.341.

Utility data and estimation of quality-adjusted life years

In order to express outcomes in the form of QALYs, the health states of the economic model need to be linked to appropriate utility scores. Utility scores represent the HRQoL associated with specific health states on a scale from 0 (death) to 1 (perfect health); they are estimated using preference-based measures that capture people's preferences on the HRQoL experienced in the health states under consideration. Preference-based measures are instruments consisting of a health state classification system, that is, an instrument that allows determination of the health state of the respondent, and an algorithm that links every health state described by the instrument with a utility score. Utility scores can also be estimated using vignettes that describe hypothetical health states including symptoms, functioning, side effects from treatment, and so on. Utility scores (which express preferences) can be elicited from various population groups (for example, service users, their parents and carers, healthcare professionals or members of the general population). The main methods of valuation are the Visual Analogue Scale (VAS), the Time Trade-Off (TTO) and the Standard Gamble (SG) (Brazier et al., 2007).

The systematic search of the literature identified 3 studies that reported utility scores associated with distinct health states experienced by adults with bipolar disorder (Depp, 2006; Hayhurst, 2006; Revicki et al., 2005).

Depp and colleagues (2006) reported utility data generated using responses to the Quality of Well-Being Scale (QWB) (Kaplan & Anderson, 1988) derived from 50 community-dwelling adults with bipolar I disorder (according to DSM-IV) aged 45 years or older; of these, 14 were in a depressive episode at the time of the evaluation, 11 in a hypomanic or manic episode, 13 in a mixed episode and 12 were in full or partial remission. The QWB scores were converted into utility scores using an algorithm that has been generated by eliciting preferences from 866 community members in the US using VAS (Kaplan & Anderson, 1988).

Hayhurst and colleagues (2006) reported EQ-5D utility values for bipolar disorder-related health states derived from 204 people with bipolar disorder participating in a multi-centre, pragmatic RCT of CBT [SCOTT2006]; participants had been recently or were still in an acute episode. The definition of health states was based on Longitudinal Interval Follow-up Evaluation (LIFE-II) Depression and Mania ratings on a 6-point scale (from l = no symptoms to 6 = DSM-IV major depressive episode, or mania with psychotic symptoms or severe impairment of function). Participants scoring 1 on both LIFE scales were considered to be in a euthymic state; those with a score of 1 or 2 on one LIFE scale and 2 on the other were considered to have residual symptoms. Adults with a score of 3 or 4 on LIFE Depression and 1 on LIFE Mania were categorised as having subsyndromal depression; those with a score of 5 or 6 on LIFE

Depression and 1 on LIFE Mania were diagnosed as depressed. No hypomanic or manic subgroup was identified within the study sample (there were only two instances of a LIFE Mania score of 5 or 6). The utility values were generated using participant responses on EQ-5D. The algorithm linking EQ-5D data to utility values has been developed following a valuation survey of 3,337 members of the general UK population using TTO (Dolan, 1997; Dolan et al., 1996).

Revicki and colleagues (2005) reported utility values of various hypothetical bipolar disorder-related health states, elicited from 96 clinically stable outpatients with bipolar I disorder in the US, using SG (values elicited using VAS were also reported). Fifty-five hypothetical health states (vignettes) were constructed for this purpose, based on reviews of psychiatric literature and consultation with psychiatrists experienced in treating bipolar disorder. Each health state described bipolar symptom severity, functioning and well-being, as well as side effects related to treatment. The study provided utility values for stable state, inpatient mania, outpatient mania and severe depression, varying with respect to the kind of pharmacological treatment obtained in each vignette and the presence or absence of side effects.

Table 5 summarises the methods used to derive and value health states associated with bipolar disorder and the resulting utility scores, as reported in the 3 studies identified in the systematic literature search conducted for this guideline.

According to NICE guidance on the selection of utility values for use in cost-utility analysis, the measurement of changes in HRQoL should be reported directly from people with the condition examined, and the valuation of health states should be based on public preferences elicited using a choice-based method, such as the TTO or SG, in a representative sample of the UK population. When changes in HRQoL cannot be obtained directly by the people with the condition examined, then data should be obtained from their carers. NICE recommends EQ-5D (Dolan, 1997) for use in cost-utility analyses of interventions for adults. When EQ-5D scores are not available or are inappropriate for the condition or effects of treatment, the institute recommends that the valuation methods be fully described and comparable to those used for the EQ-5D (NICE, 2013).

Of the three utility studies, only the one by Hayhurst and colleagues (2006) reported utility data for bipolar disorder-related health states based on EQ-5D and therefore complied with the NICE criteria on selection of appropriate utility data. However, the study reported utility values relating to depressive health states only; no relevant data on manic states were available. The study by Revicki and colleagues (2005) reported utility data associated with various bipolar disorder-related health states, including mania, acute depression and stable state. These data referred to hypothetical health states (vignettes) and were elicited from service users in the US rather than the general population,

using SG, and therefore did not satisfy NICE criteria. Finally, the study by Depp and colleagues (2006), which generated utility data from QWB scores that have been valued by members of the US general population also do not meet NICE criteria.

The GDG reviewed the available utility data against the NICE criteria, considered the limitations of each study and decided to use data from the study by Hayhurst and colleagues (2006) where possible. The reported utility value for euthymia was used for people fully responding to treatment in the economic model; the reported utility value for subsyndromal depression was used for people partially responding; and the reported utility value for depression was used for all people at the start of the model and for people not responding to treatment or relapsing to acute depression in the economic analysis.

The GDG decided to use relevant utility data from Revicki and colleagues (2005) for people relapsing to mania, due to lack of any other relevant and more appropriate data. It was decided to use for this purpose the utility value reported for inpatient mania in the study. However, the GDG noted that there were discrepancies between the values reported in Hayhurst and colleagues (2006) and Revicki and colleagues (2005) corresponding to similar health states, likely attributable to differences in the methods used by each study. For example, Revicki and colleagues (2005) reported a utility of 0.80 for the current (apparently stable) state of study participants with SG and a value of 0.67 when EQ-5D was used. The mean utility value reported for the hypothetical stable state was 0.70, that is, 0.20 lower that the respective utility value reported in Hayhurst and colleagues (2006). In addition, Revicki and colleagues (2005) reported a utility value of 0.29 for severe depression, again, almost 0.20 lower than the utility value reported for depression in the study by Hayhurst and colleagues (2006). From the above examples it can be concluded that participants in the study by Revicki and colleagues (2005) systematically under-reported the utility of bipolar disorder health states compared with participants in the study by Hayhurst and colleagues (2006). It was thus decided to add this difference of 0.20 to the utility value reported in Revicki and colleagues for inpatient mania, in order to utilise this value in the economic model.

It was assumed that all improvements and decrements in utility occurred linearly over the time period of the change in utility.

Side effects from medication are expected to result in a reduction in utility scores of adults with bipolar disorder. Disutility due to side effects was not considered in the analysis, as the model structure did not incorporate side effects. This was due to inconsistent reporting of specific side effect rates across the studies included in the network meta-analysis. This is acknowledged as a limitation of the analysis.

Table 5. Summary of studies reporting utility scores for health states experienced by adults with bipolar disorder

Study	Definition of health states	Valuation method	Population valuing	Health states and corresponding utility scores	
(Depp, 2006)	QWB data on 50 community-dwelling adults aged 45 years or older with bipolar I disorder (diagnosis based on DSM-IV)	VAS	866 community members in the US	All (n = 50) Mania or hypomania (n = 11) Mixed episode (n = 13) Depression (n = 14) Remission (n = 12)	0.54 (sd 0.09) 0.53 (sd 0.11) 0.52 (sd 0.08) 0.52 (sd 0.08) 0.59 (sd 0.10)
(Hayhurst, 2006)	EQ-5D data on 204 adults with bipolar disorder recently or still in episode participating in a multi-centre, pragmatic RCT of CBT [SCOTT2006] Definition of health states: based on LIFE-II ratings of Depression and Mania, using a 6 point scale (from l = no symptoms to 6 = DSM-IV major depressive episode or mania with psychotic symptoms or severe impairment of function). Euthymic: score = 1 on both LIFE scales Residual Symptoms: score = 1or2 on one LIFE scale and 2 on the other Subsyndromal Depression: score = 3 or 4 on LIFE Depression; 1 on LIFE Mania	TTO	3,337 members of the general UK population	Euthymic (n = 76) Residual symptoms (n = 55) Subsyndromal depression (n = 40) Depression (n = 33)	0.90 (sd 0.16) 0.83 (sd 0.16) 0.76 (sd 0.21) 0.47 (sd 0.30)

	Depressed: score = 5 or 6 on LIFE Depression; 1 on LIFE Mania			
(Revicki et al., 2005)	Hypothetical health state descriptions (vignettes) constructed based on reviews of psychiatric literature and consultation with psychiatrists experienced in treating bipolar disorder.	SG	96 clinically stable adult outpatients with DSM-IV bipolar I disorder	Current state 0.80 (sd 0.22) Stable state – no weight gain: mean (95% CI) Lithium 0.71 (0.56 to 0.86) Valproate 0.74 (0.58 to 0.89) Risperidone 0.83 (0.74 to 0.91) Olanzapine 0.82 (0.72 to 0.92) Lithium & haloperidol 0.61 (0.45 to 0.78) Valproate & haloperidol 0.62 (0.46 to 0.78) MS & risperidone 0.70 (0.62 to 0.79) MS & olanzapine 0.58 (0.48 to 0.68) MS & haloperidol 0.62 (0.51 to 0.72) No medication 0.74 (0.63 to 0.85) Stable, no medication, tardive dyskinesia 0.76 (0.64 to 0.88) Disutility because of weight gain -0.066 Severe depression 0.29 (0.16 to 0.42) Mild symptoms/SE Mean (95% CI) / Moderate symptoms/SE Mean (95% CI) Inpatient mania 0.26 (0.19 to 0.34) / 0.23 (0.16 to 0.31)

				Outpatient		
				mania	0.56 (0.39 to	0.54 (0.42 to 0.65)
				Lithium	0.73)	0.44 (0.27 to 0.62)
				Valproate	0.47 (0.30 to	0.52 (0.40 to 0.63)
				Risperidone	0.63)	0.53 (0.40 to 0.66)
				Olanzapine	0.54 (0.40 to	
				Lithium &	0.67)	0.44 (0.32 to 0.56)
				haloperidol	0.64 (0.52 to	
				Valproate &	0.76)	0.29 (0.13 to 0.44)
				haloperidol		0.41 (0.31 to 0.51)
				MS &	0.37 (0.25 to	0.53 (0.44 to 0.63)
				risperidone	0.48)	0.37 (0.28 to 0.46)
				MS &		
				olanzapine	0.63 (0.48 to	
				MS &	0.78)	
				haloperidol	0.54 (0.45 to	
					0.65)	
					0.56 (0.48 to	
					0.66)	
					0.49 (0.39 to	
					0.60)	

1 MS = mood stabiliser; TTO = Time Trade-Off; SE = side effects; SG = Standard Gamble; VAS = Visual Analogue Scale

Cost data

Costs considered in the economic model consisted of drug acquisition costs, laboratory testing costs, healthcare professional visit costs, and costs of hospitalisation and IHTTs incurred by a proportion of people not responding to treatment. Costs associated with the management of manic or depressive relapses were not considered, because these were expected to be incurred beyond the time horizon of the analysis (that is, the model was constructed in such a way that the time horizon expanded up to the point where a relapse might occur). This was decided because treatment of relapses requires a minimum of 6 to 7 weeks, and if the model was extended to include this period, people in other pathways who responded to treatment early (at 6 weeks) would be starting maintenance treatment, introducing inconsistency across different parts of the model. Costs were calculated by combining resource use estimates with respective national unit costs.

The mean daily dosage of each drug that was used in the model matched the average dosage for this drug of those reported in the relevant RCTs included in the guideline network meta-analysis, and was within the optimal dosage range according to the GDG expert opinion. Drug acquisition costs were taken from www.medicijnkosten.nl, the April 2014. For each drug the lowest reported price was selected and used in the analysis; where available, costs of generic forms were considered. Initial treatment with drugs was estimated to last 6 weeks, while people responding to treatment were assumed to receive the drug until the end of the time horizon of the analysis, that is, for 18 weeks in total, at the same daily dosage. The drug acquisition cost for no pharmacological treatment (placebo) was zero. Details on the total drug acquisition costs associated with pharmacological interventions for the treatment of acute depression in adults with bipolar disorder that were included in the economic analysis are presented in Table 6.

Table 6. Average daily dosage, acquisition costs, and 6-week and 18-week drug costs of pharmacological interventions for the management of acute depression in adults with bipolar disorder included in the economic model (2014 prices)

Drug	Mean daily dosage	Drug acquisition cost*	Total drug cost	
			6 weeks	18 weeks
Imipramine	175mg	7 x 25mg; €0.30	€12.60	€37.80
Lamotrigine	200mg	1 x 200mg; €0.13	€5.46	€16.38
Lithium	1000mg	1 x 200mg; €0.09 2 x 400mg; €0.03	€5.04	€15.12
Moclobemide	600mg	2 x 300mg; €0.74	€31.08	€93.24
Olanzapine	10mg	1 x 10mg; €0.04	€1.52	€4.56
Paroxetine	30mg	1 x 30mg; €0.04	€1.66	€4.98
Quetiapine	50% 300mg/	1 x 300mg; €0.08/ 2 x 300mg; €0.15	€4.73	€14.19

	50% 600mg			
Valproate semisodium	2000mg	4 x 500mg; €0.76	€31.85	€95.55
Fluoxetine and olanzapine	40mg and 10mg	2 x 20mg; €0.06 1 x 10mg; €0.04	€4.07	€12.20

*www.medicijnkosten.nl

People moving from first to second drug treatment following failure of first drug treatment (discontinuation or non-response) were assumed to receive the first drug at gradually reduced dosages (50% of the full dosage) for another 2 weeks following discontinuation or non-response, while the second drug was started at gradually increasing dosages (50% of the full dosage) over this 2-week period.

People moving to no pharmacological treatment following discontinuation of first drug were assumed to reduce the dosage of the discontinued drug gradually over a period of 4 weeks (each week they received 80%, 60%, 40% and 20% of the full drug dosage).

Regarding laboratory tests, according to the GDG expert opinion all cohorts in the model (including the cohort initiated on placebo) should undergo a number of tests at baseline, regardless of the initiated drug; these tests include (in Dutch):Hb, Ht, Leukocyten, differentiatie, trombocyten, Na, K, ALAT, ASAT, gammaGT, glucose, triglyceriden, cholesterol, ldl, hdl, uerum and kreatinine. There are also a number of other tests that need to be undertaken over the 18-week time horizon of the analysis that are specific to each drug. Costs were based on maximum tariffs reported by the Dutch Healthcare Authority (NZa). All laboratory tests considered in the analysis together with their unit costs are presented in Table 7.

Table 7. Laboratory tests and associated unit costs required for each pharmacological intervention received over 18 weeks for the treatment of depression in adults with bipolar disorder in the economic analysis (2014 prices)

Drug	Laboratory testing over 18 weeks	Unit costs*
Imipramine	Baseline: general tests	General tests: Hb / Ht; €1.71 Leukocyten; €1.73 Differentiatie; €1.70 Trombocyten; €1.71 Na; €1.77 K; €1.77 AF; €1.95 ALAT; €2.09 ASAT; €1.93 GammaGT; €1.93 Glucose; €1.77 Triglyceriden; €2.88 Cholesterol; €1.99
Lamotrigine	Baseline: general tests plus lamotrigine level	
Lithium	Baseline: general tests plus TSH and calcium At 12 weeks: lithium level	
Moclobemide	Baseline: general tests	
Olanzapine	Baseline: general tests At 6 weeks: glucose, cholesterol, hdl, ldl, triglyceriden At 12 weeks: nuchter glucose, cholesterol, hdl, ldl, triglyceriden	
Paroxetine	Baseline: general tests	
Quetiapine	Baseline: general tests	

Valproate semisodium	Baseline: general tests plus valproate level At 12 weeks: Hb, Ht, leucocyten, differentiatie, trombocyten, AF, ALAT, ASAT, gammaGT	LDL; €3.18 HDL; €3.18 Ureum; €1.61 Kreatinine; €1.77 **Drug specific tests:** TSH: €6.69 Calcium: €1.93 Lithium level: €4.82 Valproate level: €9.30 Lamotrigine level: €9.30
Fluoxetine and olanzapine	Baseline: general tests At 6 weeks: glucose, cholesterol, hdl, ldl, triglyceriden At 12 weeks: glucose, cholesterol, hdl, ldl, triglyceriden	

* Based on Maximum Tariffs Laboratory Research (NZa)

All people in the model received care from psychiatrists, psychologists and nurses, including those receiving no pharmacological treatment (placebo). All cohorts were assumed to have 8 contacts over the period of 18 weeks. Cohorts receiving lithium had one extra contact. In addition, people not responding to treatment or responding only partially had one additional contact. The unit cost of a contact was taken from the Handleiding voor kostenonderzoek 2010 (Hakkaart van Roijen, 2010). The mean total cost of regular contacts over 18 weeks for people responding to treatment (8 visits) was €1,583.

A proportion of people with bipolar disorder in acute depression are treated in hospital or by IHTTs. Hospitalisation and IHTT treatment rates relate to the severity of the acute episode, lack of response to treatment, and the risk of suicide and are independent of specific drug use. IHTTs are considered as an alternative to hospitalisation. According to the GDG expert opinion, the rate of hospitalisation / IHTT treatment is approximately 10% in this population. Based on data reported by Glover and colleagues (2006), it was estimated that the ratio of people with acute bipolar depression who are treated in hospital to those that are managed by CRHTTs is 77:23.

The GDG estimated that the probability of hospitalisation/IHTT management is twice as much in people who don't respond to their first drug treatment (including those who discontinued treatment) compared with those who do. Based on these estimates and the mean number of people responding to first treatment among all cohorts receiving pharmacological treatment in the model it was possible to estimate the percentage of people that are hospitalised or managed by IHTTs among those responding and those not responding to treatment, using the formulae:

$$ProbH\text{-}nr = 2 \times ProbH\text{-}r$$

$$Prob\text{-}r \times ProbH\text{-}r + Prob\text{-}nr \times ProbH\text{-}nr = ProbH$$

$$Prob\text{-}r = (1 - ProbD) \times ProbCR$$

where ProbH-nr the probability of hospitalisation/IHTT management in non-responders to first treatment (including those who discontinue their first

treatment); ProbH-r the probability of hospitalisation/IHTT management in responders to first treatment, ProbH the probability of hospitalisation/IHTT management in the total study population of adults with acute bipolar depression, estimated at 0.10, Prob-r the mean probability of response to first treatment across all cohorts in the model receiving pharmacological treatment (averaged across drug treatment options); Prob-nr the mean probability of non-response to first treatment across all cohorts, including people who discontinued treatment; and ProbD and ProbCR the mean probabilities of discontinuation conditional response, respectively, across all cohorts receiving their first pharmacological treatment, as estimated from the network meta-analysis.

Based on the above, it was estimated that the probability of hospitalisation/IHTT management in those responding to treatment was 0.063, and in those not responding was 0.126. Every person in the model was allowed to have only one incident of hospitalisation/IHTT treatment over the time horizon of the analysis.

The mean length of hospitalisation (7 weeks) was taken from data reported in the Hospital Episode Statistics for England in 2012 (NHS, 2012). Management by IHTTs was also estimated to occur over 7 weeks, according to GDG expert opinion. This was broadly consistent with the duration of CRHTT management in a RCT comparing CRHTT with standard care (inpatient services and CMHTs) for people in a psychiatric crisis in the UK (Johnson et al., 2005). People managed by IHTT in the model had 2 to 3 contacts per week, according to GDG expert opinion. The unit cost per day of hospitalisation and per IHTT contact was based on data reported in (Hakkaart van Roijen, 2010). Based on these data, the total hospitalisation cost over 7 weeks was €24,801 and the total IHTT cost was €4,048.

Costs of treating side effects of drugs were not considered in the economic analysis, due to lack of consistency in reported appropriate side effect data across all drugs. Nevertheless, the model did consider the implications of discontinuation, which is partly caused by the development of intolerable side effects. Moreover, it was estimated that the costs associated with management of side effects over the 18-week time horizon of the model were not substantial as most side effects could be dealt with during the planned contacts with the health services.

All costs have been expressed in 2014 prices, uplifted, where required, using the Consumer Price Index (www.cbs.nl). The inflation index for year 2014 was estimated using the average value of the Consumer Prices indices of the previous 4 years. As the time horizon of the analysis was less than 1 year, no discounting of costs and outcomes was necessary.

Table 12 reports the values of all input parameters utilised in the economic model and provides information on the distributions assigned to specific parameters in probabilistic analysis, as described in the next section.

Handling uncertainty

Model input parameters were synthesised in a probabilistic analysis. This means that the input parameters were assigned probabilistic distributions (rather than being expressed as point estimates), to reflect the uncertainty characterising the available clinical and cost data. Subsequently, 10,000 iterations were performed, each drawing random values out of the distributions fitted onto the model input parameters. Results (mean costs and QALYs for each intervention) were averaged across the 10,000 iterations. This exercise provides more accurate estimates than those derived from a deterministic analysis (which utilises the mean value of each input parameter ignoring any uncertainty around the mean), by capturing the non-linearity characterising the economic model structure (Briggs et al., 2006).

The distributions of the probability of discontinuation and conditional response for all pharmacological treatments as well as the probability of response for no pharmacological treatment were obtained from the network meta-analysis, defined directly from values recorded in each of the 10,000 respective iterations performed in WinBUGS. All other probabilities utilised in the economic model were given a beta distribution based on available data in the published sources of evidence and other assumptions. Utility values were also given a beta distribution using the method of moments on data reported in the relevant literature.

Drug acquisition and laboratory testing costs were not given a probabilistic distribution as these costs are set. Uncertainty in costs associated with regular and IHTT contacts was taken into account by assigning different probabilities to the number of contacts, based on expert opinion. Unit costs of regular contacts, IHTT and hospitalisation were assigned a normal distribution.

Table 8 provides details on the types of distributions assigned to each input parameter and the methods employed to define their range.

Tabel 8. Input parameters and utility data used to populate the economic model of pharmacological interventions for acute depression in adults with bipolar disorder

Input parameter	Mean value	Probabilistic distribution	Source of data - comments
Clinical input parameters	See table 4	Distribution based on network meta-analysis	Guideline network meta-analysis; distribution formed by 10,000 iterations
Probability of discontinuation, all pharmacological treatments			
Probability of conditional response, all pharmacological treatments	See table 4	Distribution based on network meta-analysis	Guideline network meta-analysis; distribution formed by 10,000 iterations
Probability of response, no pharmacological treatment (placebo)	0.35	Network meta-analysis 95% CI: 0.16 to 0.57	Guideline network meta-analysis
Ratio of probability of response: second / first line of treatment, all interventions	0.59 = 0.284/0.484	Beta distributions α = 408, β = 1031 / α = 1776, β = 1895	Rush et al., 2006
Probability of moving to no drug following discontinuation	0.25	α = 25, β = 75	GDG expert opinion; distribution based on assumption
Probability of moving to no drug following no response	0.10	α = 10, β = 90	GDG expert opinion; distribution based on assumption
Probability of partial response in responders	0.44	α = 72, β = 93	Sachs et al., 2007
3-month probability of relapse in full responders	0.08	α = 16, β = 184	Judd et al., 2008; time-dependent probabilities for each model pathway estimated from these data assuming exponential increase over time
3-month probability of relapse in partial responders	0.20	α = 40, β = 160	
Probability of mania in those relapsing	0.34	α = 43, β = 83	Judd et al., 2008
Utility values	0.47	**Beta distributions**	Hayhurst et al., 2006; distribution estimated using method of moments
Depression (baseline, no response, depressive relapse)	0.90	α = 16, β = 17	
Full response - euthymia	0.76	α = 68, β = 8	
Partial response - sub depression	0.44	α = 30, β = 10	

Mania (weighted)		α = 54, β = 69	Revicki et al. 2005, adjusted (see text for details); distribution estimated using method of moments
Resource use and costs Drug acquisition costs Laboratory testing costs	See table 6 See table 7	No distributions assigned	www.medicijnkosten.nl; Dutch Healthcare Authority (NZa)
Number of regular contacts	8	Probabilities assigned to number of contacts	GDG expert opinion; distribution based on assumption
All pathways (including placebo)	1	70%: 8; 15%: 9; 15%: 7	
Extra visits: non-responders and partial responders	0,5	70%: 1; 15%: 2; 15%: 0	
Extra visits: lithium		70%: 0,5; 25%: 1; 5%: 0	
GDG expert opinion; distribution based on assumption	17-18	50%: 17-18; 40%: 18-24; 10%: 11-17	GDG expert opinion; distribution based on assumption
Unit cost of regular contacts (2014)	€190	Normal distribution	Hakkaart van Roijen, 2010;; unit cost of regular contacts based on 'ambulant contact tweede lijn'; unit cost per hospital day based on 'algemeen ziekenhuis', unit cost per IHTT contact based on 'ambulant contact derde lijn'.distributions based on assumption
Unit cost per hospital day (2014)	€484	mean = 190, SE = 38.03	
Unit cost per IHTT contact (2014)	€192	mean = 484, SE = 24.18	
		mean = 192, SE = 9.62	
Probability of hospitalisation/IHTT	0.10	Beta distribution	GDG expert opinion; distribution based on assumption
		α = 10, β = 90	

Probability of hospitalisation/IHTT in responders	0.063	Determined by other distributions	Depending on distributions of probability of hospitalisation/IHTT, and of discontinuation and conditional response (see text for details)
Probability of hospital/IHTT in non-responders	0.126		
Proportion of IHTT in hospitalisation/IHTT	0.23	Beta distribution	Glover et al., 2006
Duration of hospitalisation/IHTT (weeks	7	$\alpha = 23, \beta = 77$	GDG expert opinion
		No distribution	

1

2

A number of deterministic one-way sensitivity analyses were undertaken to explore the impact of alternative hypotheses on the results. The following scenarios were explored:

- A change in the probability of moving to no drug following discontinuation of, or no response to, the first drug treatment option (values tested 0-1)
- A change in the probability of responsiveness to a drug if this used as second option (values tested ranged from 20% to 100% of respective probability if the drug was used as first choice)
- A change in the probability of partial response (values tested 0-1)
- A change in the probability of relapse following full or partial response (values tested 0.01-0.40 for a 3-month probability of relapse)
- A change in the overall probability of hospitalisation/IHTT management in the study population (values tested 0.02-0.20)
- An increase in the duration of hospitalisation/IHTT (values tested 8-13 weeks)
- A change in the probability of mania in case of relapse (value tested 0,25)

Presentation of the results

Results of the economic analysis are presented as follows:

For each intervention mean total costs and QALYs are presented, averaged across 10,000 iterations of the model. An incremental analysis is provided, where all options have been ranked from the most to the least effective (in terms of QALYs gained). Options that are dominated by absolute dominance (that is, they are less effective and more costly than one or more other options) or by extended dominance (that is, they are less effective and more costly than a linear combination of two alternative options) are excluded from further analysis. Subsequently, incremental cost-effectiveness ratios (ICERs) are calculated for all pairs of consecutive options remaining in analysis.

ICERs are calculated by the following formula:

$$ICER = \Delta C / \Delta E$$

where ΔC is the difference in total costs between two interventions and ΔE the difference in their effectiveness (QALYs). ICERs express the extra cost per extra unit of benefit (that is, QALY in this analysis) associated with one treatment option relative to its comparator. The treatment option with the lowest ICER is the most cost-effective option.

In addition to ICERs, the mean net monetary benefit (NMB) of each intervention is presented. This is defined by the following formula:

$$NMB = E \cdot \lambda - C$$

where E and C are the effectiveness (number of QALYs) and costs associated with the treatment option, respectively, and λ is the level of the willingness-to-

pay per unit of effectiveness, set at the cost effectiveness threshold of £20,000/QALY. The intervention with the highest NMB is the most cost-effective option (Fenwick et al., 2001). Moreover, for the most cost-effective intervention, the probability that this is the most cost-effective option is also provided, calculated as the proportion of iterations (out of the 10,000 iterations run) in which the intervention had the highest NMB among all interventions considered in the analysis.

Validation of the economic model

The economic model (including the conceptual model and the excel spreadsheet) was developed by the health economist working on this guideline and checked by a second modeller not working on the guideline. The model was tested for logical consistency by setting input parameters to null and extreme values and examining whether results changed in the expected direction. The results were discussed with the GDG for their plausibility.

Economic modelling results

The results of the economic analysis are provided in Table 9. This table provides mean QALYs and total costs for each intervention assessed in the economic analysis, as well as costs for each cost element considered in the model. Results are presented per 1000 adults with bipolar disorder in an acute depressive episode. Table 14 presents the results of the incremental analysis, the NMB of each intervention and its ranking by cost effectiveness (with higher NMBs indicating higher cost effectiveness). Interventions have been ordered from the most to the least effective in terms of number of QALYs gained.

Table 9. Results of economic analysis of pharmacological treatments for the management of acute depression in adults with bipolar disorder: mean total QALYs, total costs and detailed costs for each cost element considered in the analysis per 1000 people

Intervention	Total QALYs	Total drug cost	Total lab cost	Total Regular treatment cost	Total hospital / CRHTT cost	Total cost
Imipramine	213.76	€ 29,835	€ 54,942	€ 1,661,667	€ 1,978,151	€ 3,724,595
Lamotrigine	216.37	€ 18,546	€ 62,789	€ 1,658,417	€ 1,933,559	€ 3,673,312
Lithium	217.89	€ 17,712	€ 82,577	€ 1,764,376	€ 1,911,391	€ 3,776,056
Moclobemide	208.48	€ 55,123	€ 57,250	€ 1,667,269	€ 2,063,323	€ 3,842,965
Olanzapine	218.18	€ 11,596	€ 62,099	€ 1,656,431	€ 1,904,238	€ 3,634,363
Paroxetine	215.75	€ 12,248	€ 53,930	€ 1,659,101	€ 1,942,622	€ 3,667,901

Quetiapine	221.86	€ 16,854	€ 51,502	€ 1,652,787	€ 1,851,773	€ 3,572,916
Valproate	229.19	€ 71,780	€ 65,597	€ 1,644,739	€ 1,734,945	€ 3,517,061
Fluoxetine and olanzapine	225.81	€ 16,146	€ 62,798	€ 1,649,069	€ 1,785,527	€ 3,513,539
Placebo	198.46	€0	€0	€ 1,676,355	€ 2,006,298	€ 3,682,851

1

Table 10. Results of economic analysis of pharmacological treatments for the management of acute depression in adults with bipolar disorder: incremental analysis.

Intervention	Mean QALYs	Mean total costs	Incremental analysis and ICERs (€/QALY)		Mean NMB per person	Ranking by highest NMB
	Per 1000 people					
Valproate	229.19	€ 3,517,061	€ 1,042		€1,067	1
Fluoxetine and olanzapine	225.81	€ 3,513,539			€1,003	2
Quetiapine	221.86	€ 3,572,916	Dominated		€864	3
Olanzapine	218.18	€ 3,634,363	Dominated		€729	4
Lithium	217.89	€ 3,776,056	Dominated		€582	7
Lamotrigine	216.37	€ 3,673,312	Dominated		€654	5
Paroxetine	215.75	€ 3,667,901	Dominated		€647	6
Imipramine	213.76	€ 3,724,595	Dominated		€551	8
Moclobemide	208.48	€ 3,842,965	Dominated		€327	9
Placebo	198.46	€ 3,682,851	Dominated		€286	10

2

Valproate appears to be the most effective and cost-effective intervention, as it produces the highest number of QALYs and the highest NMB. The combination of fluoxetine and olanzapine is the next (2nd) most effective and cost-effective intervention. It is also the least costly treatment option. The ICER of valproate versus fluoxetine and olanzapine combination is €1,042/QALY, which is far below the commonly used threshold of €20,000/QALY. All other interventions are dominated by the combination of fluoxetine and olanzapine (that is, they are less effective and more costly). Quetiapine is the 3rd most cost-effective option, followed by olanzapine (4th) and lamotrigine (5th). These are followed by paroxetine (6th) and lithium (7th). imipramine and moclobemide are ranked 8th and 9th, respectively, in terms of cost effectiveness. No pharmacological treatment (placebo) is the least cost-effective intervention, ranked 10th.

The probability of valproate being the most cost-effective intervention is 0.54, which reflects the proportion of the 10,000 iterations of the economic model in which the intervention had the highest NMB among all treatment options assessed in the model. The probability of fluoxetine and olanzapine combination being the most cost-effective intervention among those assessed is 0.35. If valproate is not a treatment option, then the probability of fluoxetine and olanzapine combination being the most cost-effective intervention becomes 0.71.

Figure 2 provides the cost effectiveness plane of the analysis. Each intervention is placed on the plane according to its incremental costs and QALYs compared with placebo (which is placed at the origin).

Results were overall robust to alternative scenarios explored in sensitivity analysis. The five most cost-effective treatment options (valproate, combination of fluoxetine and olanzapine, quetiapine, olanzapine and lamotrigine) remained in the group of the five most cost-effective options in all but one of the scenarios explored. In the scenario of low hospitalization rates (0,02), paroxetine (instead of lamotrigine) ranked fifth in terms of cost-effectiveness. In some scenarios moclobemide became less cost-effective than placebo. Overall, conclusions from the analysis were not affected by the scenarios tested.

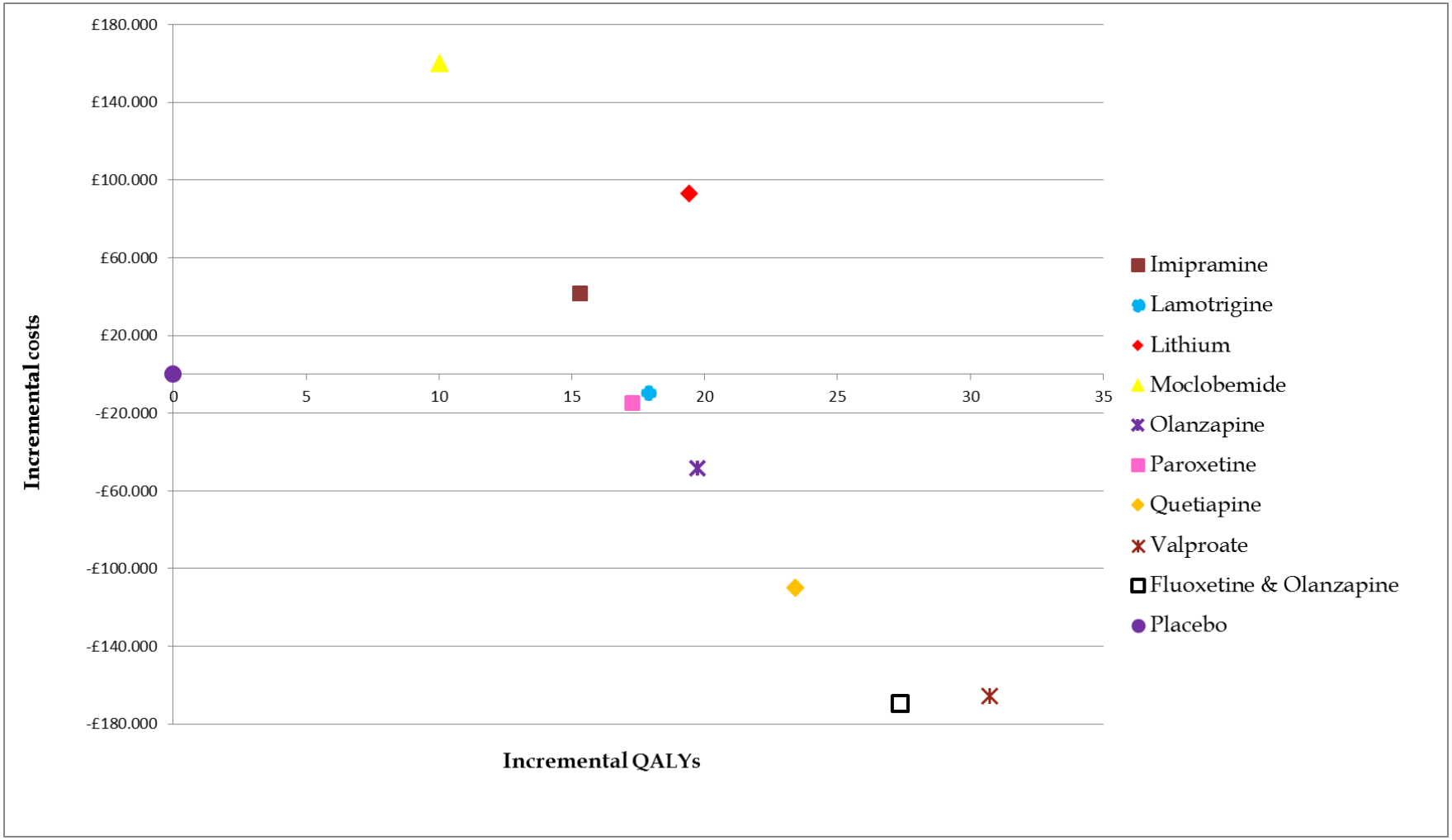
The methodology checklist and the economic evidence profile of the analysis are provided in Appendix 31 and Appendix 33, respectively.

Discussion – limitations of the analysis

The guideline economic analysis assessed the cost effectiveness of a range of pharmacological interventions for the treatment of acute depression in adults with bipolar disorder. The results of the analysis suggest that valproate may be the most cost-effective option, followed by the combination of fluoxetine and olanzapine, quetiapine, olanzapine and lamotrigine. Lithium and antidepressants used as monotherapy (paroxetine, imipramine and moclobemide) appear to be less cost-effective. These findings were not unexpected, given that the network meta-analysis did not show a statistical difference from placebo, in terms of overall response (that is, response in all randomised), for either lithium or any of the antidepressants used as monotherapy. Results were overall robust to different scenarios explored through sensitivity analysis. It should be noted that, as reported in section 1.3.4, clinical data for valproate were derived from a small number of RCT participants receiving valproate (n=48) and therefore cost effectiveness findings for this drug should be interpreted with great caution.

The clinical effectiveness data utilised in the model were derived from the network meta-analysis undertaken for this guideline. This methodology enabled evidence synthesis from both direct and indirect comparisons between interventions, and allowed simultaneous inference on all treatments examined in pair-wise trial comparisons while respecting randomisation (Caldwell et al., 2005; Lu & Ades, 2004). The assumptions and any limitations of the network meta-analysis model, as well as the limitations of individual studies considered in the network meta-analysis, have unavoidably impacted on the quality of the economic model clinical input parameters. For example, both the clinical and economic results may be vulnerable to reporting and publication bias. The assumptions underlying the network meta-analysis model have been described in detail in Appendix 15; the characteristics and any limitations of the individual studies considered in the guideline network meta-analysis model have been described in 1.3.4.

Figure 2. Cost effectiveness plane of all pharmacological interventions for acute depression in adults with bipolar disorder assessed in the economic analysis plotted against no pharmacological treatment (placebo) – incremental costs and QALYs per 1,000 people.



The economic model assumed a maximum of two lines of drugs. The purpose of considering moving to a second drug treatment option was to assess the impact of each initiated drug's non-acceptability (reflected in discontinuation rates) and ineffectiveness (reflected in non-response rates) on cost effectiveness and not to assess specific drug sequences. The clinical and cost parameters for the second pharmacological treatment option were based on the mean probabilities of discontinuation, conditional response and acquisition costs of all drug treatment options considered in the analysis, except the initiated option for each cohort. Ideally, weighted average cost and clinical outcome figures should have been used, according to actual utilisation of these drugs in the treatment of acute depression in people with bipolar disorder. However, specific data on actual drug utilisation patterns for adults with acute bipolar depression were not possible to find.

There are indications that treatment with antidepressants may induce switching to mania, although this appears to be a controversial issue (Baldessarini et al., 2013; Sidor & McQueen, 2011; Tondo et al., 2010). The risk of switching to mania associated with antidepressants was not considered in the model due to lack of good quality data in the RCTs included in the guideline network meta-analysis and the wider literature. The GDG suggested that any available data on this issue be considered in a sensitivity analysis. Nevertheless, this analysis proved unnecessary as the base-case analysis demonstrated that antidepressants were not cost-effective. Consideration of switching to mania would only increase the costs for these drugs (due to high hospitalisation costs associated with mania), thus reducing their relative cost effectiveness even more.

The impact of side effects on quality of life and associated management costs was not considered in the analysis, due to lack of appropriate relevant data. However, omission of important side effects (such as the renal failure associated with lithium and the acute extrapyramidal syndrome and weight gain associated with antipsychotics) from the model structure is unlikely to have affected the results of the analysis due to its short time horizon. Moreover, some short-term side effects have been taken into account implicitly in the model structure, since discontinuation of treatment occurs to some extent due to the development of intolerable side effects. Also, a number of short-term side effects can be dealt with by routine contacts with health services at no additional cost. In addition, the probabilistic model allowed a small proportion of people to have a higher number of regular contacts, which could be relating to management of side effects.

Therefore, although omission of side effects is acknowledged as a limitation of the analysis, it is estimated that it has not impacted considerably on the results.

Some clinical input parameters were taken from studies that were not directly relevant to the model population and condition. For example, data on the

potential reduction in responsiveness following second treatment were taken from a study on people with unipolar (rather bipolar) depression (Rush et al., 2006) because of lack of more relevant data. The probability of partial response in those responding was based on relevant recovery (rather than response) data on people with bipolar depression (Sachs et al., 2007); partial recovery in that study was defined by the duration of effect, rather than its intensity. The probability of relapse following response was estimated using data on relapse after recovery (not response) from any acute major episode, not just depressive, in people with bipolar disorder (Judd et al., 2008). Some data on resource use (especially the overall probability of hospitalisation/IHTT management in the study population) were based on the GDG expert opinion, due to lack of relevant data. The impact of all these parameters was tested in sensitivity analysis, which suggested that the results were robust under a broad range of alternative values and scenarios.

Costs associated with treatment of relapses were not considered in the model, because the model was constructed in such a way that the time horizon expanded up to the point where a relapse might occur. This was decided so as to avoid introducing long-term maintenance treatment to people in some pathways in the model (which would occur if the model was extended to capture the management of relapses), and thus inconsistency in the treatment received across pathways. It should be clarified that the model did not consider the reduction in utility occurring during a manic or depressive relapse, but it did consider the deterioration in HRQoL from the point of response to treatment and up to the point of (but not including) relapse. This allowed a more realistic representation of the HRQoL during the period following response for people eventually relapsing.

Another limitation of the analysis was its short time horizon. Ideally, the analysis should consider longer-term outcomes of the acute treatment, including modelling of long-term maintenance treatment. However, this was not possible due to lack of relevant long-term data across the drugs considered in the analysis. On the other hand, the time horizon of 18 weeks was adequate as it enabled the full course of acute bipolar depression to be modelled, and the associated costs and benefits from pharmacological treatment to be assessed.