Appendix E: Network meta-analyses – methods and detailed results

Network meta-analysis (NMA) was undertaken to inform decision making for 2 review questions in this guideline – those concerning the effectiveness of PPIs for severe erosive oesophagitis (healing and maintenance phases; see full guideline section 4.4.3.1) and the effectiveness of different eradication regimens for *H pylori* (first- and second-line options; see full guideline sections 4.7.3 and 4.7.7).

This appendix explains methods used for NMAs, highlighting any deviations from the Guidelines Manual (2012), and presents full results to accompany the summary results presented in the full guideline.

E.1 Introduction

In a decision problem comparing more than 2 mutually exclusive treatment options, the results of conventional pairwise meta-analyses of direct evidence alone are unlikely fully to inform a decision about which option is most effective. The challenge of interpretation has arisen for the following reasons:

- In isolation, each direct pairwise comparison cannot fully inform the choice between all the different options; therefore, a series of discrete pairwise comparisons can be difficult to interpret.
- Invariably, direct comparisons of some treatments of interest are not available. For example, option A may be compared, in separate trials and analyses, with options B and C, but there is no direct evidence of the relative effectiveness of treatments B and C.

NMAs overcome these issues by allowing all evidence to be combined in a single, internally consistent model, synthesising data from direct and indirect comparisons whilst preserving the randomisation of the RCTs included in the reviews. The resulting syntheses produce estimates of relative effectiveness for all comparators and ranking of different interventions.

The terms indirect treatment comparisons, mixed treatment comparisons and NMA are often used interchangeably in the published literature. We use the term NMA as the networks conducted for this guideline consist of both indirect treatment comparisons (some trials have a common comparator and some do not) and mixed treatment comparisons (with at least one closed loop, combination of direct and indirect evidence).

E.2 Synthesis methods

General methods common to all NMAs undertaken for this guideline are detailed below. Any additional steps taken in approaching individual questions are discussed

E.2.1 Implementation of syntheses

We undertook hierarchical Bayesian NMA using WinBUGS version 1.4.3. The models used reflected the recommendations of the NICE Decision Support Unit's Technical Support Documents (TSDs) on evidence synthesis, particularly TSD 2 ('A generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials'; see http://www.nicedsu.org.uk/). We used the WinBUGS code provided in the appendices of TSD 2 without substantive alteration to specify synthesis models.

We report results summarising 10,000 samples from the posterior distribution of each model, having first run and discarded 50,000 ‘burn-in’ iterations. Three separate chains with different initial values were used.

E.2.2 Prior distributions

Non-informative prior distributions were used in all models. Trial-specific baselines and treatment effects were assigned $N(0, 1000)$ priors, and the between-trial standard deviations used in random-effects models were given $U(0, 5)$ priors. These are consistent with the recommendations in TSD 2 for dichotomous outcomes.

E.2.3 Dichotomous outcomes

As advised in TSD 2, dichotomous outcomes can be synthesised using 2 alternative models:

- The most straightforward model adopts a **binomial likelihood** with a **logit link function**, and generates output on a **log-odds scale**, with results transformed to odds ratios for presentation.
- An alternative model incorporates data on duration of follow-up in each underlying RCT, assuming a constant rate of events, to estimate the probability of events occurring over time. Again, a binomial likelihood is assumed, but a **complementary log–log ('cloglog') link function** is used, which results in outputs on a **log-hazard scale** (transformed into hazard ratios for presentation).

Where differences in follow-up in the underlying evidence were believed or shown to be minor and/or unimportant, the simpler logit-link model was preferred. Where duration of follow-up was believed to have a potential impact on outcomes, both models were explored, and the choice made on the basis of goodness of fit (see E.2.4).

Zero cells

In datasets containing studies with 'zero cells' (that is, trials in which no events occurred in 1 or more arm), substantial instability was encountered when performing syntheses. To address this problem, a constant of 0.5 was added to all cell counts (effectively adding 0.5 to the numerator and 1 to the denominator of the proportion). The same approach was used to address instability for datasets containing studies with 100% events reported in all arms.

Studies reporting no events in any arms were excluded from NMAs, as they do not provide any information on the relative likelihood of events occurring.

E.2.4 Choice of reference treatment

To undertake an NMA, the analyst must specify 1 treatment in the network as a common ‘reference’ option in comparison to which the model will estimate the treatment effects of all other options. The choice of reference treatment is mathematically arbitrary; however, it may have implications for the computational efficiency of the network and/or the interpretability of outputs. For these reasons, it is advisable to choose an option that is well connected within the network (that is, one that has been compared with as many of the other treatments as possible). A ‘standard treatment’ or placebo option often provides a good choice, because it will usually be well represented in the underlying evidence, and it also provides a readily understood common comparator for summary outputs (that is, everything else compared with placebo will be easier to interpret than everything compared with an option with which some readers are unfamiliar).

E.2.5 Goodness of fit

Measures of model fit were scrutinised to assess appropriateness of each model. Particular attention was paid to:

- **Total residual deviance**: a calculation of the model's ability to predict the individual datapoints underlying it. In every iteration of the model sampling procedure, the amount each model-estimated datapoint deviates from the observed evidence is calculated, summed and averaged over all iterations. Each datapoint should contribute about 1 to the posterior mean deviance; therefore, the total residual deviance of a well fitting model will be approximately the same as the number of independent datapoints in the model
- **Deviance information criterion (DIC)**: an estimate of deviance that is 'penalised' according to the number of parameters in the model (adding parameters to a model should increase its ability to predict known data; however, this may come at the expense of reducing its ability to predict external datasets).
- **SD of random-effects term (tau)**: where a random-effects model is fitted, the width of the inter-study heterogeneity distribution estimated by the model is a reflection of how well the model accounts for heterogeneity in the underlying data. Therefore, while not a measure of goodness of fit *per se*, it is useful to consider as an indication of how broad a model is required to fit the data. There is no analogous quantity for fixed-effects models.

E.2.6 Reported outputs

The NMA outputs shown in the full guideline and/or this appendix are as follows:

- Network diagram, showing availability of evidence. These diagrams have the following features:
  - The size of each node is proportional to total number of participants randomised to receive the treatment in question across the evidence-base.
  - The width of connecting lines is proportional to number of trial-level comparisons available.
  - Arrowheads indicate direction of effect in pairwise data (a > b denotes a is more effective than b) – filled arrowheads show comparisons where one option is significantly superior ($p<0.05$); outlined arrowheads show direction of trend where effect does not reach statistical significance.
- Table of input data, showing the evidence used in the model.
- Relative effect matrix, showing an estimate of effect for each regimen compared with each of its comparators; an estimate of effect based on direct evidence only (pairwise frequentist meta-analysis using fixed- or random-effects models as in the NMA) is also presented for comparisons where data is available
- Plot of the relative effectiveness, including the results of the NMA of each regimen compared with the reference treatment (see E.2.4) and any direct estimate available for the same comparison.
- Tabulated rank probabilities, giving the probability of each treatment being best (that is, ranked #1) and its median rank with 95% credible interval (CrI). In these outputs, higher ranking always reflect what is best for the patient (for example: higher rates of disease eradication, lower rates of adverse events, higher IQ, lower blood pressure, and so on).
- Histograms demonstrating the probability of each treatment being at each possible rank ('rankograms')

B.2.7 Quality assessment

B.2.7.1 GRADE for pairwise meta-analyses

GRADE was used to assess the quality of outcomes as specified in the Guidelines manual (2012).

B.2.7.2 Modified GRADE for NMAs

As there is no published guidance for using GRADE with NMAs, a modified approach was adopted:

- A quality rating was assigned, based on the study design, therefore as each NMA contained only RCTs they started at 'high'
- The rating was then downgraded for risk of bias, inconsistency, imprecision and indirectness using the criteria detailed below. Each quality element considered to have 'serious' or 'very serious' limitations was rated down 1 or 2 levels respectively.

Risk of bias

The overall quality of evidence for each outcome was considered for risk of bias and assessed conventionally for each included trial. These were then compiled as an overall assessment for the entire group of included studies within the NMA for the following criteria:

- Appropriateness of randomisation method
- Adequacy of concealment methods (blinding)
- Study design – outcomes were downgraded if the methodology used for outcome detection was not clear. For example, reporting methods for some outcomes were poor across the studies, in particular methods used to obtain data on adverse events and adherence to medication were often not reported or unclear

For this criterion it is also important to assess how the risk of bias from the direct comparisons may have an effect on the indirect comparisons within the network. Therefore, the risk of bias was assessed for each direct comparison and then an assessment was made about how the risk of bias from the direct comparisons would affect the indirect comparisons. Additionally, there was an assessment of treatment effect modifiers and if they differed between links in the network.

Inconsistency

Within a NMA inconsistency refers to unexplained heterogeneity (that is, widely differing estimates of treatment effect across studies that suggest true differences in the underlying treatment effect) between the direct and indirect comparisons (i.e. the 'loops' of data within the network diagrams). Therefore, evidence may be downgraded in quality if there is inconsistency between indirect estimates produced by the NMA and direct estimates that are obtained from pairwise comparisons in included trials. Heterogeneity across studies for each direct pairwise meta-analysis was assessed using I^2. This allowed for the assessment of heterogeneity within the included studies using the following decision rules:

- The NMA was downgraded 1 level when there was observed heterogeneity (I^2>50%) for 1 link or more in a network, but there were also links with no observed heterogeneity
- The NMA was downgraded 2 levels for inconsistency if all links within the network had considerable (I^2>50%), substantial (I^2>30%) or moderate (I^2 >10%) heterogeneity

Additionally, to assess for inconsistency for each pairwise comparison where both direct and indirect evidence were available, the values of the direct and indirect estimates were compared to see if they were similar.

Indirectness

When assessing indirectness within an NMA the quality of the evidence was not downgraded for indirectness due to the use of indirect comparisons, as this is taken into account within the other GRADE criteria. The evidence could however be downgraded in quality if an indirect population, intervention, comparator or outcome was used, as in conventional pairwise comparisons for GRADE.

Imprecision

Imprecision relates to the overall level of confidence that may be placed in the estimated treatment effects. As currently there is no guidance on how to set MIDs and no guidance on the defaults MIDs in the context of Bayesian statistics with 95% credible intervals in NMAs, evidence was downgraded if there was uncertainty around the indirect estimates and the probability ranking of relative treatments. This was judged for the following variables:

- The number of studies within each link used to form the network
- The number of direct head-to-head trials
- Event rates within the included trials
- Assessment of the CrI in terms of degree of overlapping with each other.

The number of studies within each link used to form the network, event rates and the resulting width of CrI were the main criteria considered for this review as all included trials were head-to head:

- For the purposes of this guideline when the majority of links contained only 1 trial the NMA was downgraded 1 level
- The quality of the evidence was downgraded by 1 level where the total number of events was less than 300 (a threshold rule-of-thumb value for frequentist analysis but considered to be applicable to Bayesian analysis) (based on: Mueller et al. Ann Intern Med. 2007;146:878-881)
- When overlaps of the 95% CrI occurred in >50% of the point estimates for each node the NMA was downgraded by 1 level. When the 95% CrI overlapped in >75% of the point estimates for each node the NMA was downgraded by 2 levels

E.3 PPIs for severe erosive oesophagitis

E.3.1 Question-specific methods

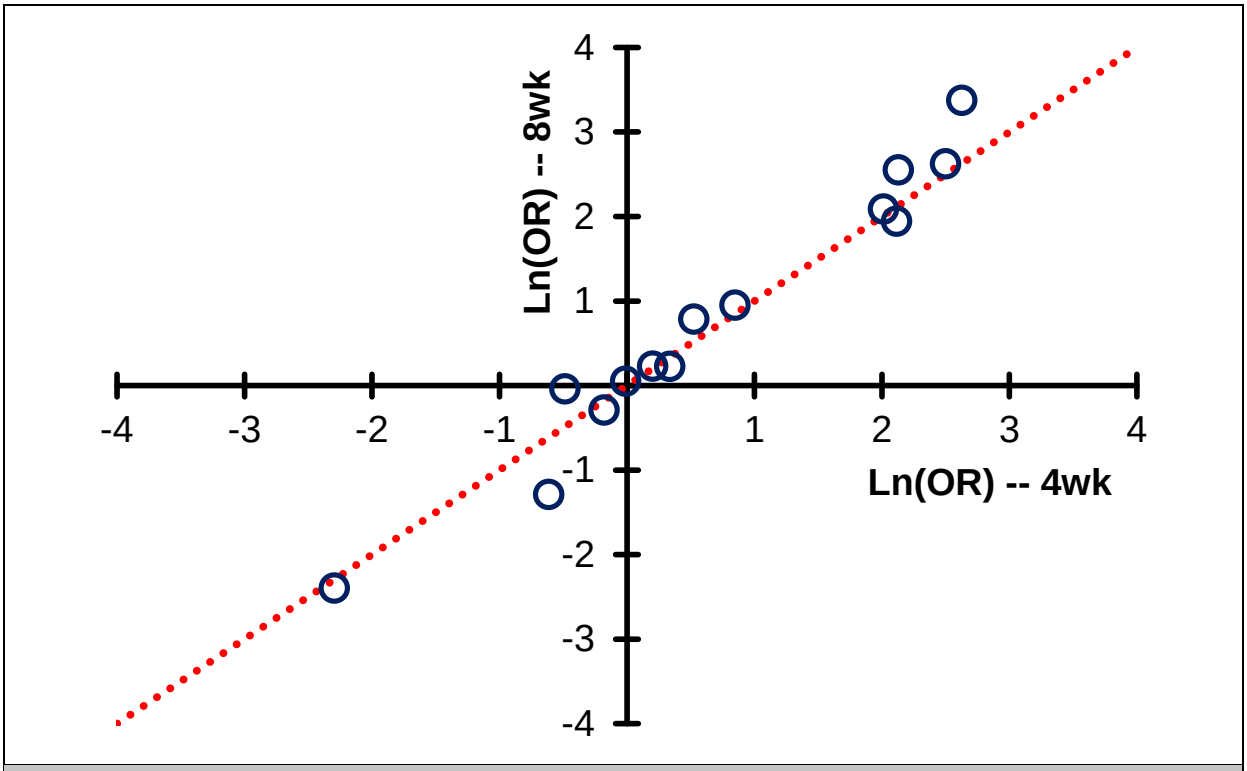
E.3.1.1 Healing

E.3.1.1.1 Selection of data

The critical outcome for this question is probability of healing, as assessed by endoscopy. Included RCTs reported these data after 4 and/or 8 weeks of treatment. It would have been possible to perform separate NMAs for each juncture; however, this would have led to sparse evidence networks, with some treatment options represented at 4 weeks' follow-up but not at 8, and vice versa. Consequently, we explored the possibility of using data from both junctures in a single synthesis.

We compared the relative effect measures from RCTs reporting at both 4 and 8 weeks and found that there was a very strong correlation between the 2 junctures (Figure 1). This means that the degree to which one treatment is better than another is very closely comparable at both timepoints (that is, if drug A is twice as good as drug B at achieving

healing after 4 weeks, it will be twice as good at 8 weeks, too, although the absolute probability of healing will rise for both options as treatment extends).



Dotted line indicates x=y

Figure 1: Comparison of relative effect (log odds ratios) for healing in trials reporting both 4- and 8-wk follow-up

Having established the equivalence of relative effect at the 2 junctures, we considered it was appropriate to pool data from both 4- and 8-week timepoints to estimate the relative effectiveness of all comparators. However, using both junctures from any individual RCT would amount to double-counting of data. Therefore, the datapoints used reflect the latest follow-up available in each RCT (that is, 4-week data are only used for RCTs that do not provide 8-week data).

E.2.1.1.2 Reference treatment

Pantoprazole 40mg/d was selected as the reference treatment, as it is connected to all other options by the fewest number of links (it is common to use placebo as a reference treatment, where available; however, it would not be sensible to do so in this instance, as the amount of placebo-controlled evidence is small and, as can be seen in Figure 2, it is peripheral to the network).

E.2.1.1.3 Models used

We used logit-link binomial models (see E.2.2). Fixed- (FE) and random-effects (RE) models were fitted for each network. The RE model was selected as the residual deviance for the RE model was closer to the number of unconstrained datapoints and DIC was lower (Table 1).

Table 1: Measures of goodness of fit of fixed- and random-effects models for the healing network

Measure of goodness of fit	FE model	RE model
Residual deviance*	53.39	43.44
Deviance information criterion (DIC)	260.863	255.766
Tau	n/a	0.002 (95%CrI: 0.004, 0.654)
*Compared to 41 datapoints		

E3.1.2 Maintenance

E3.1.2.1 Selection of data

The critical outcome is probability of relapse, as assessed by endoscopy.

The evidence network for this question presented a problem for coherent analysis, as it consisted of 2 discrete, disconnected networks (firstly, pantoprazole at 10 mg/d, 20 mg/d and 40 mg/d compared with ranitidine 300 mg/d and, secondly, lansoprazole at 15 mg/d and 30 mg/d compared with esomeprazole 20 mg/d and placebo). Analysis of these separate networks would enable inference to be drawn about the relative effectiveness of options within each group, but it would not be possible to reach conclusions about how treatments from different sub-networks compare with each other. To overcome this problem, the GDG agreed to consider pantoprazole 10 mg/d as equivalent to placebo, thereby merging the nodes and providing a common point of comparison for all treatments. The justification for this decision was twofold: firstly, the GDG noted that 10 mg/d is half the recommended minimum dose for pantoprazole (hence, it would not be expected to have more than a placebo effect in practice); secondly, inspection of the raw data supported this a priori expectation – the relapse rate in the 1 placebo arm in the evidence-base was 74% and the 2 pantoprazole 10 mg/d arms had relapse rates of 73% and 100% (see Table 7). Consequently, the GDG were happy to treat the two options as equivalent.

E3.1.2.2 Reference treatment

Once placebo and pantoprazole 10 mg/d had been combined to form a single comparator (see above), it was sensible to use this as the reference treatment for the network, both because it is central to and well connected in the evidence-base and because it makes comparisons readily interpretable.

E3.1.2.3 Models used

Included RCTs reported relapse rate after either 6 or 12 months' follow-up. In contrast to the 4- and 8-week datapoints in the healing phase evidence-base (see above), there were no trials reporting both these junctures; therefore, it was not possible to assess whether relative effects can be assumed to change as follow-up extends. For this reason, 2 different models were explored for the maintenance dataset – 1 that, in an identical way to the healing-phase NMA, combined effectiveness estimates regardless of duration of follow-up (log-odds scale; binomial likelihood; logit link function) and one that incorporated data on duration of follow-up to estimate effects on a log-hazard scale (binomial likelihood; complementary log–log ['cloglog'] link function; see E.2.2).

We fitted FE and RE versions of each model and examined measures of goodness of fit to discriminate between them (Table 2).

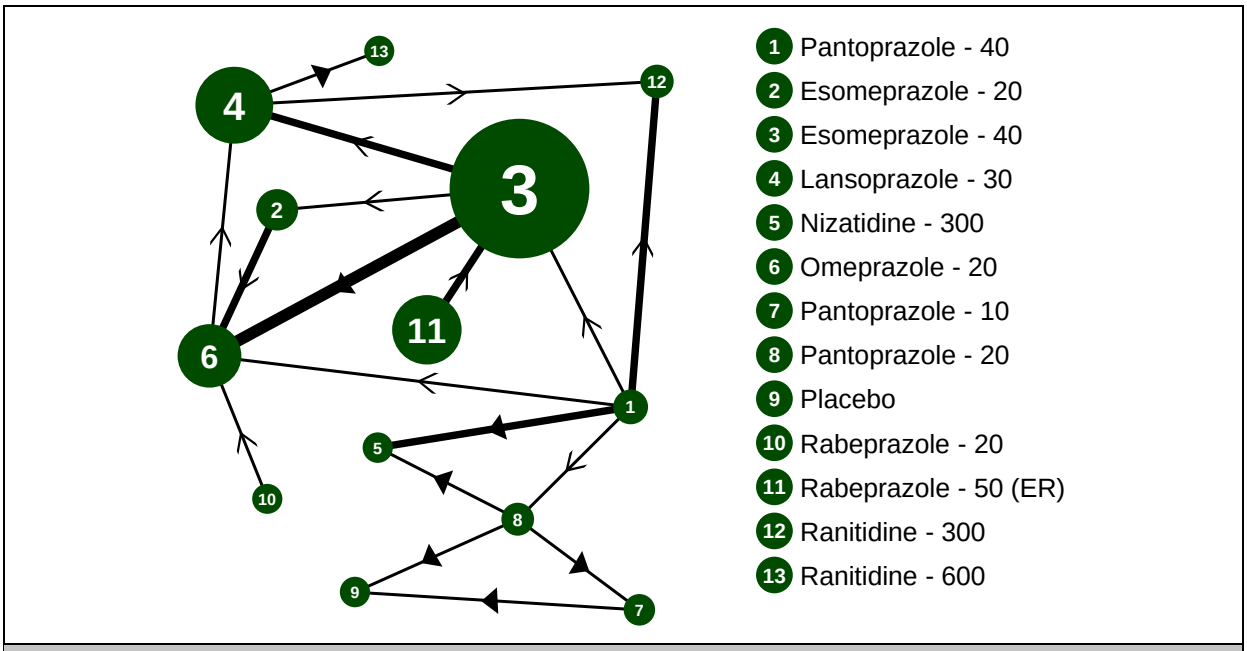
Table 2: Measures of goodness of fit for candidate models for the maintenance network

Measure of goodness of fit	Logit-link odds ratio		Cloglog-link hazard ratio	
	FE	RE	FE	RE
Residual deviance*	24.14	17.41	20.79	15.4
Deviance information criterion (DIC)	89.989	86.623	86.518	84.536
Tau	n/a	1.085 (95%CrI: 0.102, 1.943)	n/a	0.726 (95%CrI: 0.068, 1.843)
*Compared to 15 datapoints				

The RE version of the cloglog modle was found to have a superior fit to the data (as assessed by lower residual deviance and DIC), so was preferred for all analyses.

E.3.2 Results

E.3.2.1 Healing



Size of nodes is proportional to total number of participants randomised to receive the treatment in question across the evidence-base. Width of connecting lines is proportional to number of trial-level comparisons available. Arrowheads indicate direction of effect in pairwise data (a > b denotes a is more effective than b) – filled arrowheads show comparisons where one option is significantly superior (p<0.05); outlined arrowheads show direction of trend where effect does not reach statistical significance.

Figure 2: Network meta-analysis of healing (4-8wks) – evidence network

Table 3: Network meta-analysis of healing (4–8wks) – input data

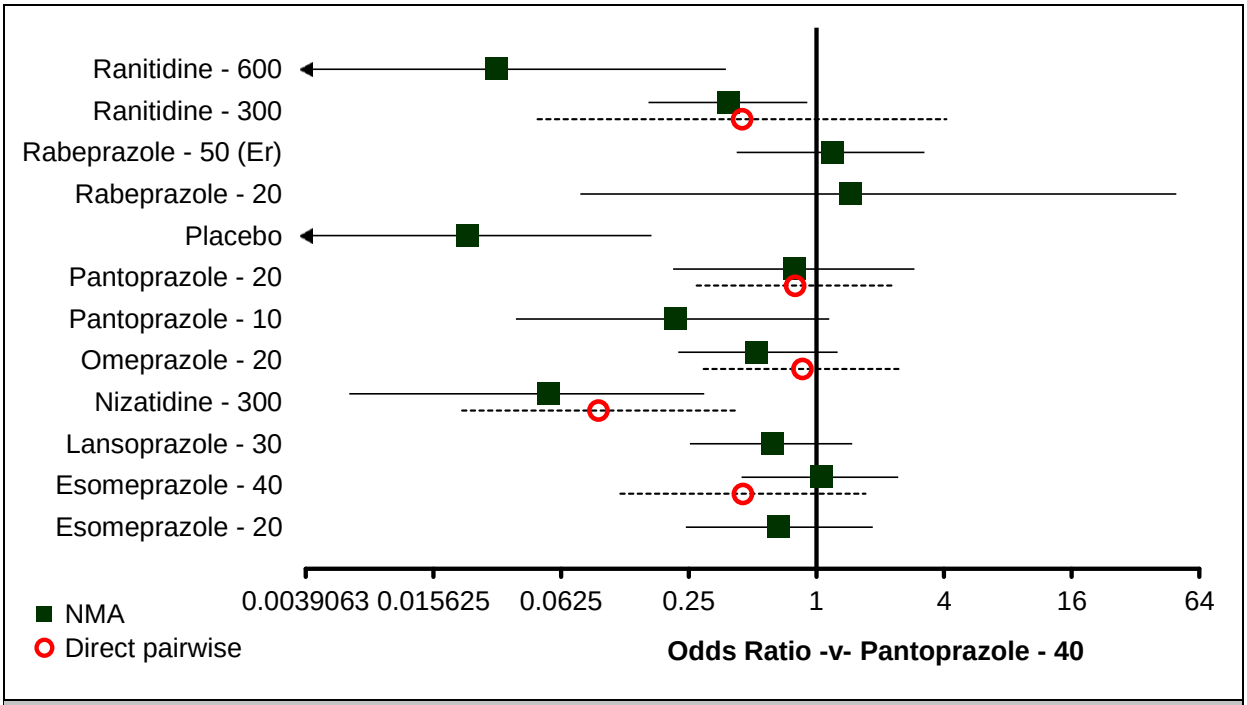
	Pantoprazole - 40	Esomeprazole – 20	Esomeprazole - 40	Lansoprazole - 30	Nizatidine - 300	Omeprazole - 20	Pantoprazole - 10	Pantoprazole - 20	Placebo	Rabeprazole - 20	Rabeprazole - 50 (ER)	Ranitidine - 300	Ranitidine - 600
Armstrong (2001) – 4wk	1/6				0/6								
Castell (2002) – 8wk			552/640	477/646									
Fennerty (2005) – 8wk[a]			386/498	367/501									
Gillessen (2004) – 10wk[b]	12/18		9/19										
Jansen (1999) – 8wk				10/11									7/16
Kahrilas (2000) – 8wk		124/165	136/166			133/182							
Koop (1995) – 4wk	17/30											9/14	
Kovacs (2002) – 8wk[a]	16/27				2/21			15/28					
Laine(a) (2011) – 8wk[a]			398/531								419/524		
Laine(b) (2011)			421/537								409/528		
Lightdale (2006) – 8wk		122/158				110/154							
Mee (1996) – 8wk[a]				26/37		27/38							
Meneghelli (2002) – 8wk[a]	20/24											10/24	
Mossner (1995) – 4wk	21/36					12/22							
Pace (2005) – 8wk						13/15				14/15			
Richter (2000) – 8wk[a]							23/60	45/65	2/28				
Richter (2001) – 8wk[a]			268/317			217/320							
Robinson (1995) – 8wk				48/63								46/71	
Schmitt (2006) – 8wk[a]			167/189			131/169							

[a] *Data also available for 4wk follow-up; only 8wk data used in analysis*
[b] *Assumed same as 8wk in analyses*

Table 4: Network meta-analysis of healing (4–8wks) – relative effectiveness of all pairwise combinations

	Pantoprazole - 40	Esomeprazole - 20	Esomeprazole - 40	Lansoprazole - 30	Nizatidine - 300	Omeprazole - 20	Pantoprazole - 10	Pantoprazole - 20	Placebo	Rabeprazole - 20	Rabeprazole - 50(ER)	Ranitidine - 300	Ranitidine - 600
Pantoprazole - 40		-	0.45 (0.12,1.70)	-	0.09 (0.02,0.41)	0.86 (0.29,2.50)	-	0.79 (0.27,2.31)	-	-	-	0.45 (0.05,4.11)	-
Esomeprazole - 20	0.67 (0.24,1.84)		1.50 (0.88,2.55)	-	-	0.82 (0.58,1.16)	-	-	-	-	-	-	-
Esomeprazole - 40	1.06 (0.44,2.43)	1.59 (0.77,3.05)		0.60 (0.34,1.04)	-	0.45 (0.34,0.59)	-	-	-	-	1.12 (0.80,1.57)	-	-
Lansoprazole - 30	0.63 (0.25,1.48)	0.93 (0.41,2.03)	0.59 (0.36,0.98)		-	1.04 (0.38,2.81)	-	-	-	-	-	0.58 (0.27,1.23)	0.08 (0.01,0.76)
Nizatidine - 300	0.05 (0.01,0.30)	0.08 (0.01,0.59)	0.05 (0.01,0.35)	0.09 (0.01,0.58)		-	-	10.96 (2.14,56.3)	-	-	-	-	-
Omeprazole - 20	0.53 (0.22,1.26)	0.79 (0.43,1.46)	0.50 (0.32,0.82)	0.84 (0.48,1.61)	9.92 (1.47,97.0)		-	-	-	2.15 (0.17,26.7)	-	-	-
Pantoprazole - 10	0.22 (0.04,1.15)	0.32 (0.04,2.20)	0.20 (0.03,1.33)	0.34 (0.05,2.27)	4.01 (0.51,48.0)	0.40 (0.06,2.62)		3.62 (1.73,7.59)	0.12 (0.03,0.57)	-	-	-	-
Pantoprazole - 20	0.79 (0.21,2.90)	1.17 (0.22,5.88)	0.75 (0.16,3.51)	1.27 (0.26,6.08)	14.38 (2.63,132)	1.50 (0.31,6.80)	3.62 (1.30,10.6)		0.03 (0.01,0.16)	-	-	-	-
Placebo	0.02 (0.00,0.17)	0.03 (0.00,0.33)	0.02 (0.00,0.19)	0.04 (0.00,0.33)	0.40 (0.03,6.55)	0.04 (0.00,0.38)	0.11 (0.01,0.51)	0.03 (0.00,0.14)		-	-	-	-
Rabeprazole - 20	1.46 (0.08,49.9)	2.13 (0.12,72.3)	1.34 (0.08,44.9)	2.28 (0.13,77.9)	28.34 (0.9,1441)	2.68 (0.16,86.5)	6.60 (0.25,358)	1.80 (0.08,83.7)	67.10 (1.8,5366)		-	-	-
Rabeprazole - 50(ER)	1.19 (0.42,3.23)	1.79 (0.71,4.29)	1.13 (0.63,2.03)	1.91 (0.88,4.14)	22.63 (3.03,232)	2.27 (1.03,4.64)	5.55 (0.77,39.7)	1.52 (0.29,7.79)	53.12 (5.40,777)	0.84 (0.02,15.5)		-	-
Ranitidine - 300	0.39 (0.16,0.91)	0.58 (0.19,1.72)	0.37 (0.15,0.94)	0.62 (0.26,1.50)	7.30 (1.06,73.9)	0.73 (0.28,1.88)	1.80 (0.28,12.4)	0.49 (0.10,2.37)	17.14 (1.89,243)	0.27 (0.01,5.14)	0.32 (0.11,0.99)		-
Ranitidine - 600	0.03 (0.00,0.37)	0.05 (0.00,0.53)	0.03 (0.00,0.32)	0.05 (0.00,0.52)	0.57 (0.01,15.2)	0.06 (0.00,0.64)	0.14 (0.00,2.96)	0.04 (0.00,0.67)	1.40 (0.02,42.8)	0.02 (0.00,0.97)	0.03 (0.00,0.30)	0.08 (0.00,0.99)	

Values given are odds ratios.
The segment below and to the left of the shaded cells is derived from the network meta-analysis, reflecting direct and indirect evidence of treatment effects (row versus column). The point estimate reflects the mean of the posterior distribution, and numbers in parentheses are 95% credible intervals. The segment above and to the right of the shaded cells gives pooled direct evidence (random-effects pairwise meta-analysis), where available (column versus row). Numbers in parentheses are 95% confidence intervals.



Values less than 1 favour Pantoprazole 40; values greater than 1 favour the comparator treatment. Solid error bars are 95% credible intervals while dashed error bars are 95% confidence interval.

Figure 3: Network meta-analysis of healing (4–8wks) – relative effect of all options compared with placebo

Table 5: Network meta-analysis of healing (4–8wks) – rankings for each comparator

	Probability best	Median rank (95%CrI)
Pantoprazole - 40	0.105	3 (1, 7)
Esomeprazole - 20	0.011	6 (2, 9)
Esomeprazole - 40	0.054	3 (1, 6)
Lansoprazole - 30	0.002	6 (3, 9)
Nizatidine - 300	0.000	11 (10, 13)
Omeprazole - 20	0.000	7 (4, 10)
Pantoprazole - 10	0.002	10 (3, 11)
Pantoprazole - 20	0.122	5 (1, 9)
Placebo	0.000	12 (11, 13)
Rabeprazole - 20	0.482	2 (1, 11)
Rabeprazole - 50 (ER)	0.221	2 (1, 7)
Ranitidine - 300	0.001	9 (4, 10)
Ranitidine - 600	0.000	12 (9, 13)

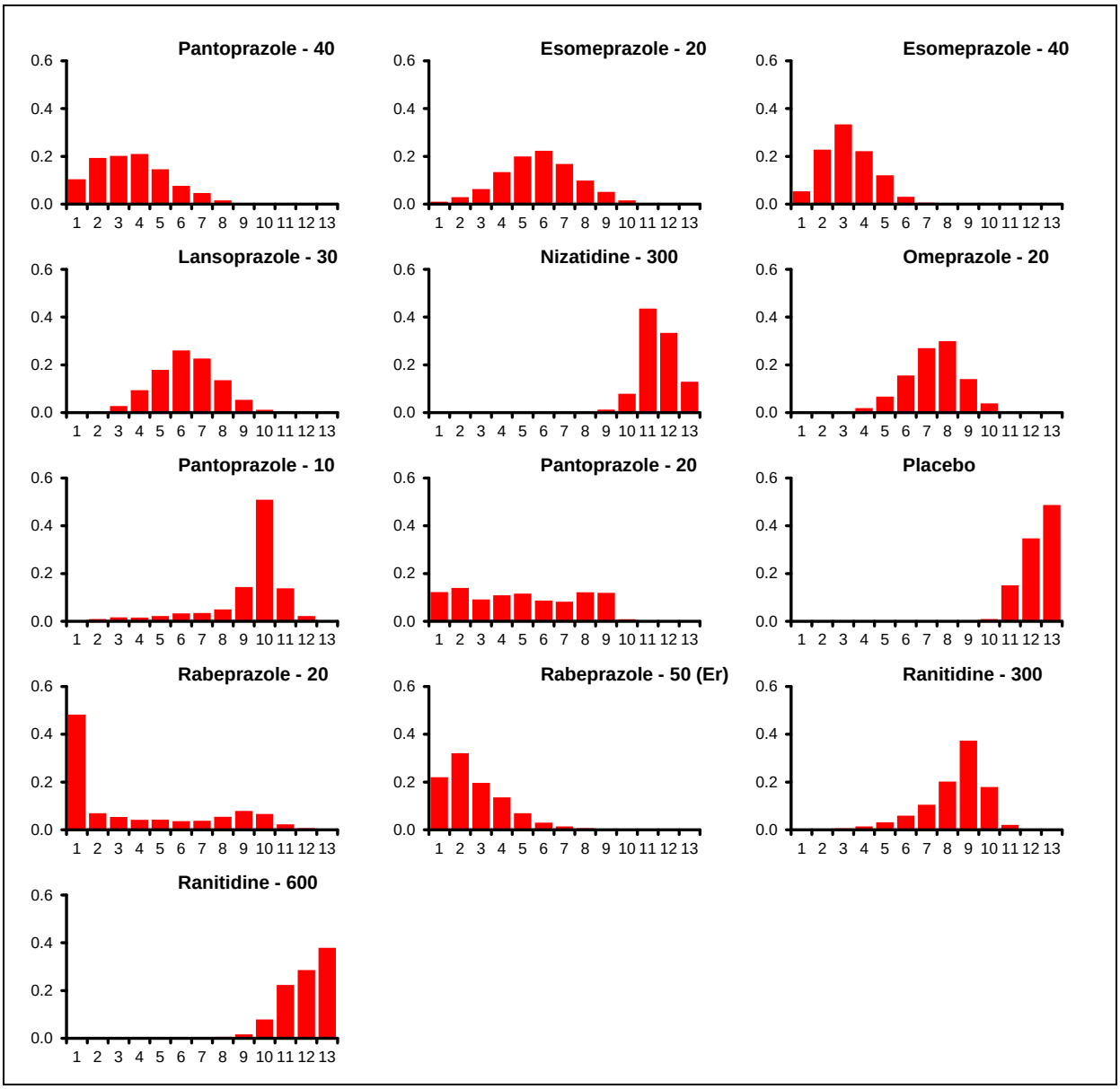
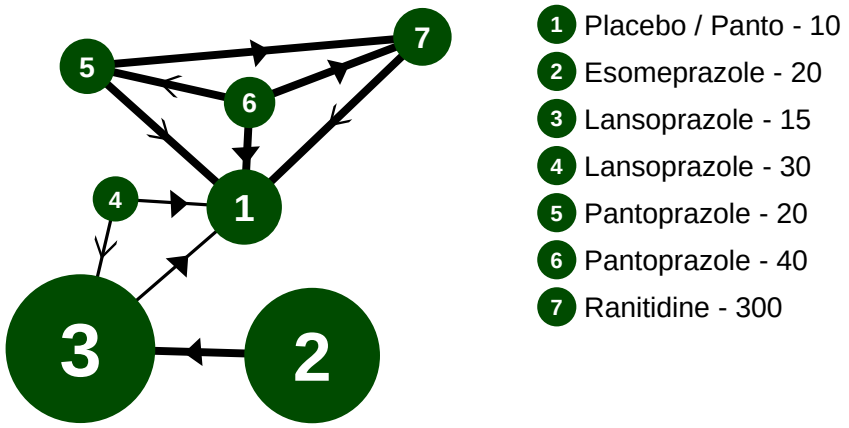


Figure 4: Network meta-analysis of healing (4–8wks) – rank probability histograms

Table 6: Network meta-analysis of healing (4–8wks) – model fit statistics

Residual deviance	Dbar	Dhat	pD	DIC	tau
43.41 (compared to 41 datapoints)	219.796	183.906	35.89	255.687	0.294 (95%CrI: 0.054, 0.793)

E.3.2.2 Maintenance



Size of nodes is proportional to total number of participants randomised to receive the treatment in question across the evidence-base. Width of connecting lines is proportional to number of trial-level comparisons available. Arrowheads indicate direction of effect in pairwise data (a > b denotes a is more effective than b) – filled arrowheads show comparisons where one option is significantly superior (p<0.05); outlined arrowheads show direction of trend where effect does not reach statistical significance.

Figure 5: Network meta-analysis of relapse (6–12mo) – evidence network

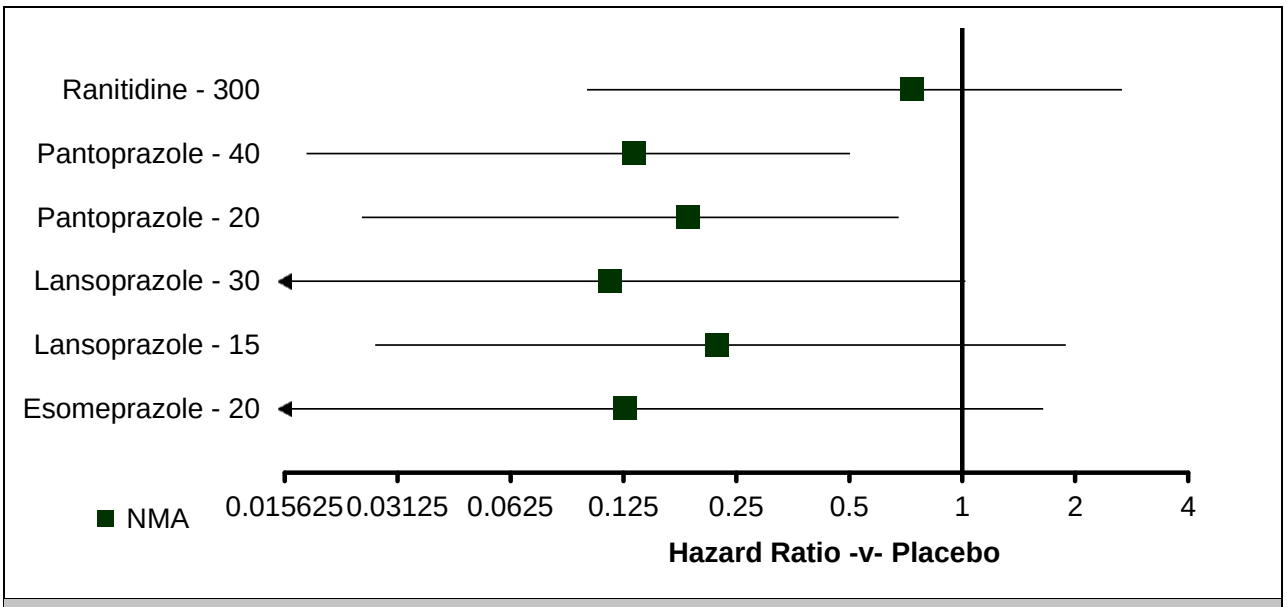
Table 7: Network meta-analysis of relapse (6–12mo) – input data

	Placebo / Panto - 10	Esomeprazole - 20	Lansoprazole - 15	Lansoprazole - 30	Pantoprazole - 20	Pantoprazole - 40	Ranitidine - 300
Robinson (1996) – 1yr	26/35		9/33	5/32			
Lauritsen (2003) – 0.5yr		27/114	42/102				
DeVault (2006) – 0.5yr		25/121	40/131				
Richter (2004) – 1yr	22/30				14/31	5/19	21/26
Metz (2003) – 1yr	34/34				8/23	10/26	31/34

Table 8: Network meta-analysis of relapse (6–12mo) – relative effectiveness of all pairwise combinations

	Placebo / Panto 10	Esomeprazole - 20	Lansoprazole - 15	Lansoprazole - 30	Pantoprazole - 20	Pantoprazole - 40	Ranitidine - 300
Placebo							
Esomeprazole - 20	0.13 (0.01, 1.64)						
Lansoprazole - 15	0.22 (0.03, 1.89)	1.76 (0.41, 7.57)					
Lansoprazole - 30	0.12 (0.01, 1.02)	0.91 (0.06, 13.14)	0.52 (0.05, 4.81)				
Pantoprazole - 20	0.19 (0.03, 0.68)	1.47 (0.05, 22.34)	0.83 (0.04, 8.63)	1.61 (0.07, 19.72)			
Pantoprazole - 40	0.13 (0.02, 0.50)	1.05 (0.04, 16.05)	0.59 (0.03, 6.38)	1.15 (0.05, 14.58)	0.72 (0.15, 3.36)		
Ranitidine - 300	0.74 (0.10, 2.67)	5.76 (0.20, 88.64)	3.27 (0.16, 33.92)	6.28 (0.30, 74.27)	3.91 (0.90, 17.84)	5.42 (1.21, 25.27)	

Values given are hazard ratios. The segment below and to the left of the shaded cells is derived from the network meta-analysis, reflecting direct and indirect evidence of treatment effects (row versus column). The point estimate reflects the mean of the posterior distribution, and numbers in parentheses are 95% credible intervals. Because it is not easily possible to derive analogous estimates of hazard ratios from a frequentist analysis of direct data only, the segment above and to the right of the shaded cells is left blank.



Values less than 1 favour the comparator treatment; values greater than 1 favour placebo. Solid error bars are 95% credible intervals.

Figure 6: Network meta-analysis of relapse (6–12mo) – relative effect of all options compared with placebo

Table 9: Network meta-analysis of relapse (6–12mo) – rankings for each comparator

	Probability best	Median rank (95%CrI)
Placebo / Pantoprazole 10	0.000	7 (5, 7)
Esomeprazole - 20	0.294	2 (1, 6)
Lansoprazole - 15	0.020	4 (2, 7)
Lansoprazole - 30	0.338	2 (1, 6)
Pantoprazole - 20	0.096	4 (1, 5)
Pantoprazole - 40	0.249	3 (1, 5)
Ranitidine - 300	0.003	6 (3, 7)

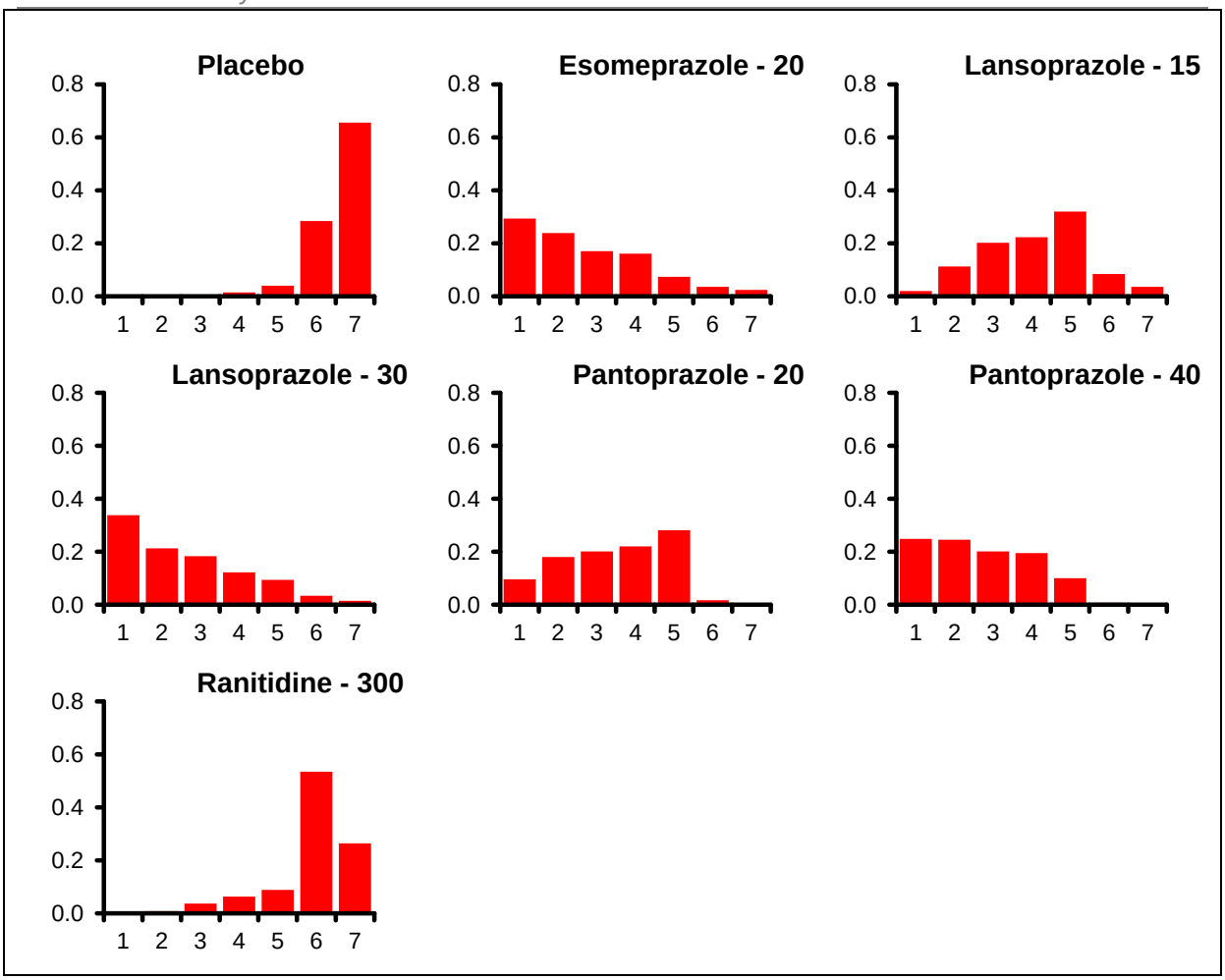


Figure 7: Network meta-analysis of relapse (6–12mo) – rank probability histograms

Table 10: Network meta-analysis of relapse (6–12month) – model fit statistics

Residual deviance	Dbar	Dhat	pD	DIC	tau
15.47 (compared to 15 datapoints)	70.247	55.828	14.42	84.667	0.712 (95%CrI: 0.055, 1.845)

E.4 *H pylori* eradication

E.4.1 Question-specific methods

E.4.1.1 Study selection and data collection

To estimate the relative efficacy of different *H pylori* eradication regimens for first and second-line treatment, NMAs were conducted using included RCT evidence identified for the review questions.

Five NMAs were conducted, defined by population and outcome measure:

First-line eradication

- Eradication network

Second-line eradication

- Eradication network
- Adverse events (rash) network
- Adverse events (loose stools) network
- Adherence to medication network

E.4.1.2 Reference treatment

We selected the following as reference treatments in the networks:

- **First-line eradication**: PPI/AMO/CLA as this regimen was recommended in the previous guideline (CG17)
- **Second-line eradication (all outcomes)**: PPI/BIS/NIT/TET as this regimen was most fully represented in the evidence-base

E.4.1.3 Models used

We explored 2 alternative models for synthesising dichotomous outcomes (see E.2.2). There were negligible differences between results from the two types of model. However, it was observed that the cloglog model can be unstable when there are no or few events in either arm (even when a constant was added to studies with zero cells); this problem was particularly common for individual adverse events. For this reason, logit models were used in the final syntheses. It was also noted that producing results as odds ratios may be more helpful for model validation, as they provide a straightforward point of comparison with frequentist syntheses of direct evidence.

Fixed versus random effects

FE and RE models were fitted for each network. The model selected for each network and rationale for selection is outlined in the tables below.

Table 11: Measures of goodness of fit of fixed- and random-effects models for the first-line eradication network

Measure of goodness of fit	FE model	RE model
Residual deviance*	59.35	44.49
Deviance information criterion (DIC)	254.885	245.905
Tau	n/a	0.630 (95%CrI: 0.232, 1.458)
*Compared to 41 datapoints		

The RE model was selected as the residual deviance for the RE model was closer to the number of unconstrained datapoints and DIC was lower.

Table 12: Measures of goodness of fit of fixed- and random-effects models for the second-line eradication network

Measure of goodness of fit	FE model	RE model
Residual deviance*	42.07	38.76

Measure of goodness of fit	FE model	RE model
Deviance information criterion (DIC)	210.781	210.626
Tau	n/a	0.678 (95%CrI: 0.045, 1.854)
*Compared to 36 datapoints		

The RE model was selected as the residual deviance for the RE model was close to the number of unconstrained datapoints, although it was noted that there was very little to choose between the models in DIC.

Table 13: Measures of goodness of fit of fixed- and random-effects models for the second-line adverse event (rash) network

Measure of goodness of fit	FE model	RE model
Residual deviance*	25.72	25.72
Deviance information criterion (DIC)	93.959	94.611
Tau	n/a	0.851 (95%CrI: 0.046, 1.931)
*Compared to 24 datapoints		

There was only a marginal difference in the residual deviance for both models and they were both relatively close to the number of unconstrained datapoints. The FE model was selected due to its slight advantage in DIC and more parsimonious interpretation.

Table 14: Measures of goodness of fit of fixed- and random-effects models for the second-line adverse event (loose stools) network

Measure of goodness of fit	FE model	RE model
Residual deviance*	34	33.47
Deviance information criterion (DIC)	143.053	143.832
Tau	n/a	0.710 (95%CrI: 0.036, 1.872)
*Compared to 32 datapoints		

There was only a marginal difference in the residual deviance for both models and they were both relatively close to the number of unconstrained datapoints. The FE model was selected due to its slight advantage in DIC and more parsimonious interpretation.

Table 15: Measures of goodness of fit of fixed- and random-effects models for the second-line adherence to medication network

Measure of goodness of fit	FE model	RE model
Residual deviance*	21.63	22.06
Deviance information criterion (DIC)	105.973	107.06
Tau	n/a	0.670 (95%CrI: 0.031, 1.892)
*Compared to 22 datapoints		

Both models had residual deviance that was very close to the number of unconstrained datapoints. The FE model was selected due to its slight advantage in DIC and more parsimonious interpretation.

8.4.2 Results

8.4.2.1 First-line eradication

Eradication network

A total of 16 RCTs from the evidence review met the inclusion criteria for the eradication NMA. Six studies which reported eradication could not be included in the NMA for the following reasons:

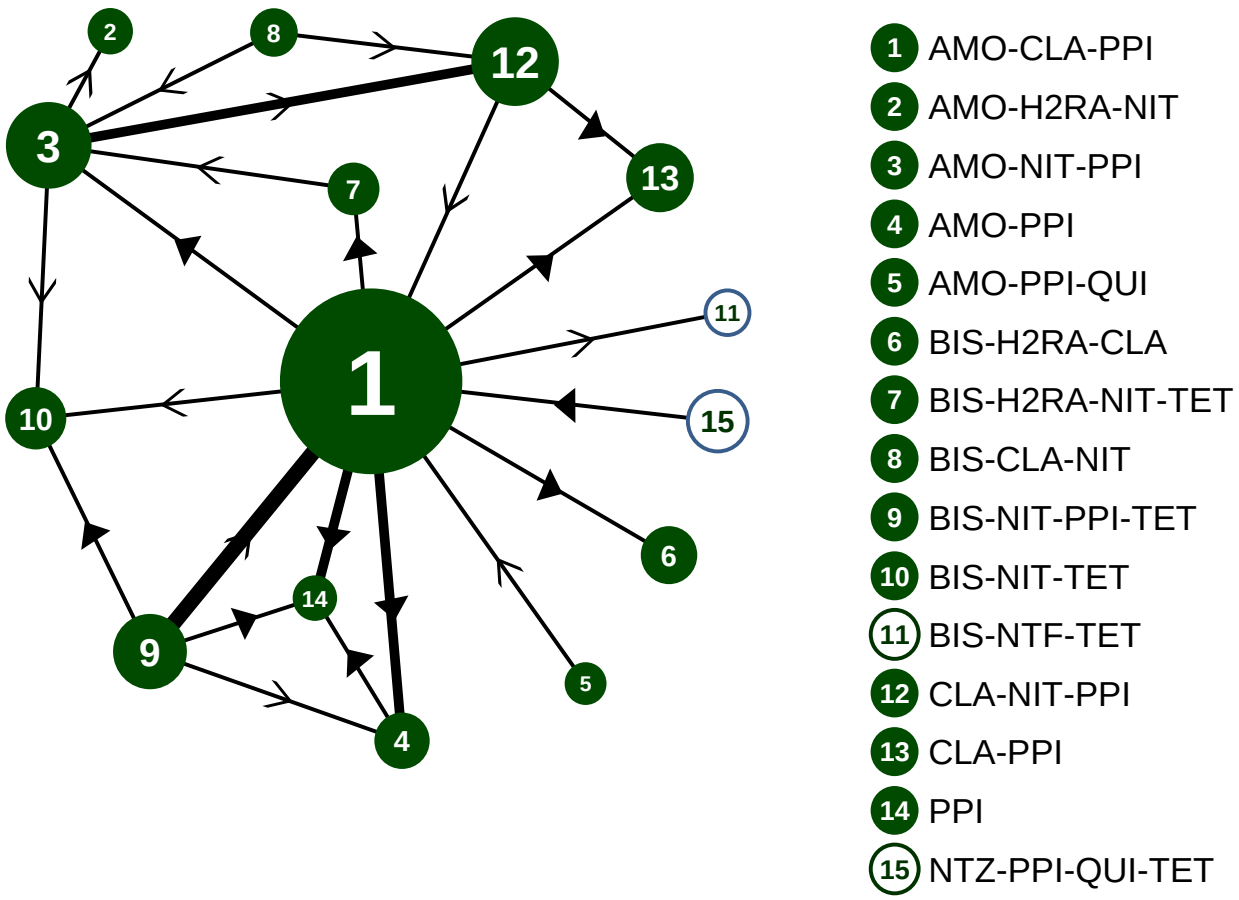
- The studies compared the same regimens and only the dose or duration differed (internal loop: denoted by a dashed line on the network) (5 studies)
- One comparison was not linked to the network and therefore could not be compared (indirectly) with the regimens in the network (1 study)

The data from these studies was analysed using pairwise meta-analysis and was presented to the GDG and considered alongside the outputs from the NMA. Full GRADE tables can be viewed in appendix F.

Table 17 shows the odds ratio matrix for first-line *H pylori* eradication and summarises the results of the conventional pairwise meta-analyses together with the results generated by the NMA for every possible treatment comparison. The section below and to the left of the shaded diagonal is derived from the NMA, reflecting the combined direct and indirect evidence of treatment effects (row versus column). The point estimate reflects the mean of the posterior distribution, and numbers in parentheses are 95%CrI. The section above and to the right of the shaded diagonal gives pooled direct evidence (pairwise meta-analysis) where available (column versus row). Numbers in parentheses are 95% CI.

The plot of relative effectiveness (Figure 9) shows indirect estimates of interventions in comparison to standard first-line treatment (PPI/AMO/CLA) from the NMA with 95% CrI (solid error bars) and direct pairwise head-to-head comparisons with 95% CI (dashed error bars) in graphical form.

The rank probability table (Table 18) and rankograms (Figure 10 & Figure 11) show the probability of ranking in each position for each regimen for achieving *H pylori* eradication. Results are given separately for the network including and excluding regimens including an unlicensed component; this is because the inclusion of options that could only be recommended in unusual circumstances may conceal differences between other options.



Outlined nodes with dark numbers represent regimens with an unlicensed component.
Size of nodes is proportional to total number of participants randomised to receive the treatment in question across the evidence-base. Width of connecting lines is proportional to number of trial-level comparisons available.
Arrowheads indicate direction of effect in pairwise data (a > b denotes a is more effective than b) – filled arrowheads show comparisons where one option is significantly superior (p<0.05); outlined arrowheads show direction of trend where effect does not reach statistical significance.

Figure 8: Eradication - evidence network

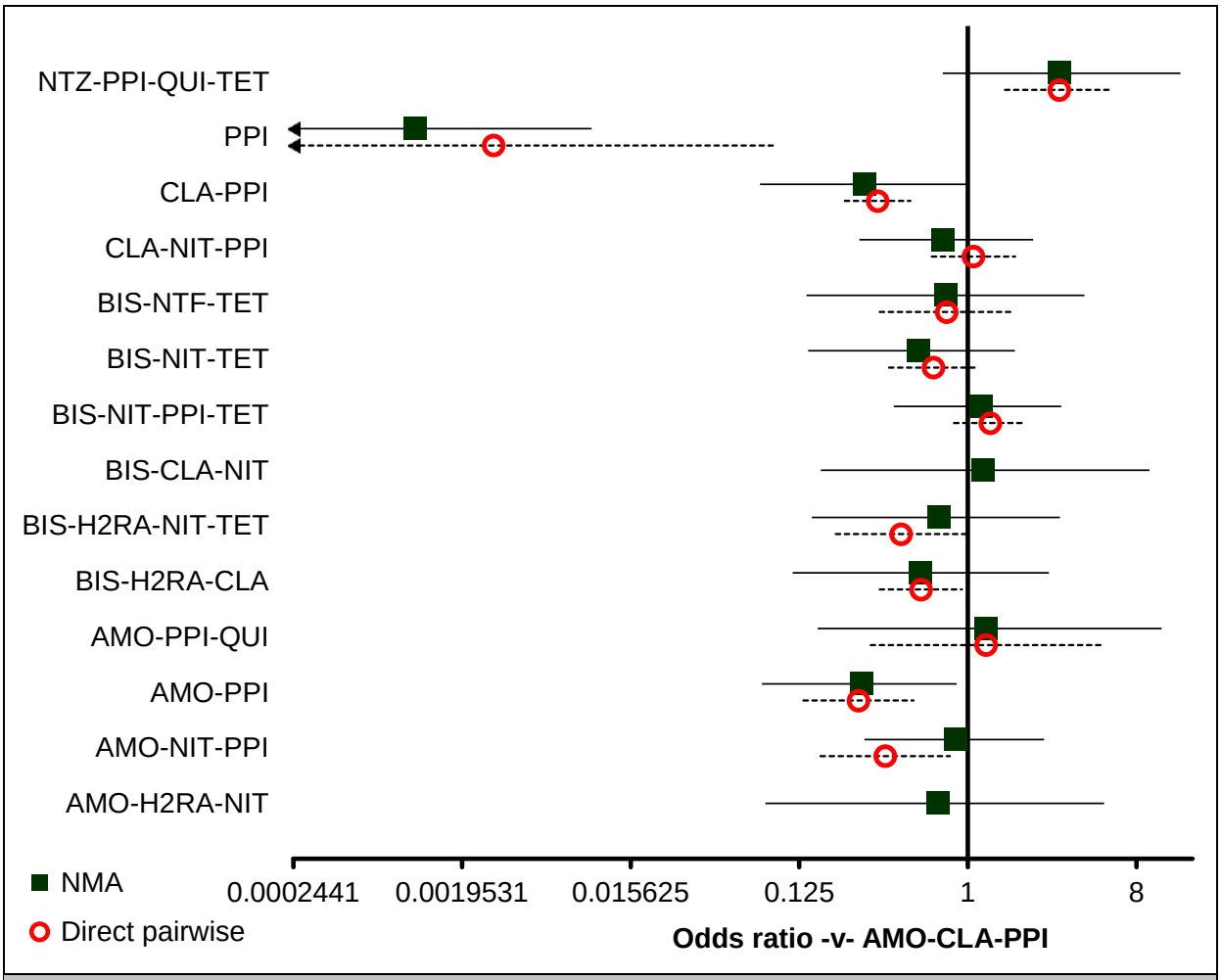
Table 16: Network meta-analysis of first-line eradication of *H pylori* – input data

	AMO-CLA-PPI	AMO-H_2RA-NIT	AMO-NIT-PPI	AMO-PPI	AMO-PPI-QUI	BIS-H_2RA-CLA	BIS-H_2RA-NIT-TET	BIS-CLA-NIT	BIS-NIT-PPI-TET	BIS-NIT-TET	BIS-NTF-TET	CLA-NIT-PPI	CLA-PPI	PPI	NTZ-PPI-QUI-TET
Antos et al. (2006)	26/31				26/30										
Arkkila et al. (2005)	27/27			24/29					25/27					0/29	
Basu et al. (2011)	66/90														161/180
Chiba (1996)												28/34	18/31		
Ecclissato et al. (2002)	27/46										24/46				
Hsu et al. (2001)		48/60	50/60												
Katelaris et al. (2000)			89/109									64/111			
Katelaris et al. (2002)	104/134								110/134	95/137					
Koivisto et al. (2005)	100/110		83/106				92/113								
Laine et al. (2000) Trial A	179/233												112/215		
Laine et al. (2000) Trial B	58/74													1/24	
Laine et al. (2003)	114/137								121/138						
Lee et al. (1999)	83/116											140/192			
Lerang et al. (1997)			44/46							49/54					
Lerang et al. (1997)			70/77					74/78				72/76			
Ohlin et al. (2002)	48/62			56/115											
Veldhuyzen van Zanten et al. (2003)	118/152					101/153									

Table 17: Network meta-analysis of first-line eradication of *H pylori* – relative effectiveness of all pairwise combinations

	AMO-CLA-PPI	AMO-H_2RA-NIT	AMO-NIT-PPI	AMO-PPI	AMO-PPI-QUI	BIS-H_2RA-CLA	BIS-H_2RA-NIT-TET	BIS-CLA-NIT	BIS-NIT-PPI-TET	BIS-NIT-TET	BIS-NTF-TET	CLA-NIT-PPI	CLA-PPI	PPI	NTZ-PPI-QUI-TET
AMO-CLA-PPI		-	0.36 (0.16,0.80)	0.26 (0.13,0.51)	1.25 (0.30,5.19)	0.56 (0.34,0.93)	0.44 (0.20,0.98)	-	1.32 (0.84,2.05)	0.65 (0.38,1.13)	0.77 (0.34,1.75)	1.07 (0.64,1.79)	0.33 (0.22,0.49)	0.00 (0.00,0.10)	3.08 (1.58,6.00)
AMO-H_2RA-NIT	0.70 (0.08,5.37)		1.25 (0.49,3.16)	-	-	-	-	-	-	-	-	-	-	-	-
AMO-NIT-PPI	0.87 (0.28,2.56)	1.25 (0.21,7.50)		-	-	-	1.21 (0.63,2.35)	1.85 (0.52,6.60)	-	0.45 (0.08,2.41)	-	0.68 (0.12,3.81)	-	-	-
AMO-PPI	0.27 (0.08,0.87)	0.39 (0.04,4.35)	0.31 (0.06,1.57)		-	-	-	-	2.60 (0.46,14.7)	-	-	-	-	0.00 (0.00,0.07)	-
AMO-PPI-QUI	1.25 (0.16,10.9)	1.80 (0.10,37.0)	1.44 (0.14,16.4)	4.62 (0.42,56. 9)		-	-	-	-	-	-	-	-	-	-
BIS-H_2RA-CLA	0.56 (0.12,2.72)	0.81 (0.06,11.6)	0.65 (0.10,4.55)	2.07 (0.30,15.4)	0.45 (0.03,5.99)		-	-	-	-	-	-	-	-	-
BIS-H_2RA-NIT-TET	0.70 (0.15,3.11)	1.00 (0.10,10.4)	0.81 (0.18,3.69)	2.59 (0.36,18.6)	0.56 (0.04,7.29)	1.26 (0.13,10.9)		-	-	-	-	-	-	-	-
BIS-CLA-NIT	1.21 (0.16,9.41)	1.74 (0.14,23.0)	1.38 (0.23,9.40)	4.45 (0.44,48.5)	0.97 (0.05,17.4)	2.18 (0.17,29.4)	1.71 (0.17,19.0)		-	-	-	0.97 (0.23,4.04)	-	-	-
BIS-NIT-PPI-TET	1.18 (0.40,3.17)	1.71 (0.17,16.3)	1.36 (0.31,5.61)	4.37 (1.00,18.2)	0.95 (0.08,9.22)	2.12 (0.30,12.9)	1.69 (0.27,9.99)	0.97 (0.10,8.65)		0.49 (0.28,0.87)	-	-	-	0.00 (0.00,0.04)	-
BIS-NIT-TET	0.55 (0.14,1.78)	0.79 (0.08,7.45)	0.63 (0.14,2.53)	2.01 (0.35,10.4)	0.43 (0.03,4.52)	0.98 (0.12,6.64)	0.78 (0.11,4.83)	0.45 (0.04,3.93)	0.46 (0.12,1.73)		-	-	-	-	-
BIS-NTF-TET	0.76 (0.14,4.22)	1.10 (0.08,16.7)	0.87 (0.12,6.77)	2.80 (0.36,22.8)	0.60 (0.04,9.24)	1.35 (0.13,14.4)	1.08 (0.11,11.5)	0.63 (0.04,8.91)	0.64 (0.09,4.91)	1.40 (0.17,12.6)		-	-	-	-
CLA-NIT-PPI	0.74 (0.26,2.24)	1.07 (0.14,8.80)	0.85 (0.31,2.56)	2.72 (0.58,14.2)	0.59 (0.05,6.15)	1.32 (0.19,9.31)	1.05 (0.21,6.03)	0.62 (0.09,3.95)	0.62 (0.16,2.87)	1.34 (0.33,6.86)	0.97 (0.13,7.60)		0.30 (0.10,0.92)	-	-
CLA-PPI	0.28 (0.08,0.99)	0.41 (0.04,4.34)	0.32 (0.07,1.55)	1.04 (0.18,6.03)	0.22 (0.02,2.52)	0.51 (0.07,3.87)	0.40 (0.06,2.84)	0.23 (0.02,2.08)	0.24 (0.05,1.26)	0.52 (0.10,3.22)	0.37 (0.04,3.10)	0.38 (0.10,1.44)		-	-
PPI	0.00 (0.00,0.01)	0.00 (0.00,0.03)	0.00 (0.00,0.01)	0.00 (0.00,0.04)	0.00 (0.00,0.02)	0.00 (0.00,0.03)	0.00 (0.00,0.02)	0.00 (0.00,0.02)	0.00 (0.00,0.01)	0.00 (0.00,0.02)	0.00 (0.00,0.02)	0.00 (0.00,0.02)	0.00 (0.00,0.05)		-
NTZ-PPI-QUI-TET	3.11 (0.73,13.7)	4.48 (0.36,60.2)	3.58 (0.59,23.2)	11.45 (1.84,76.9)	2.49 (0.19,32.6)	5.56 (0.66,47.8)	4.40 (0.54,38.7)	2.59 (0.21,30.9)	2.62 (0.48,16.6)	5.67 (0.92,44.0)	4.10 (0.43,37.9)	4.21 (0.68,25.5)	10.96 (1.63,79.5)	2892.00 (213,123900)	

Values given are odds ratios.
The segment below and to the left of the green diagonal cells is derived from the network meta-analysis, reflecting direct and indirect evidence of treatment effects (row versus column). The point estimate reflects the mean of the posterior distribution, and numbers in parentheses are 95% credible intervals. The segment above and to the right of the green diagonal cells gives pooled direct evidence (random-effects pairwise meta-analysis), where available (column versus row). Numbers in parentheses are 95% confidence intervals.



Values less than 1 favour AMO-CLA-PPI; values greater than 1 favour the comparator treatment. Solid error bars are 95% credible intervals while dashed error bars are 95% confidence interval.

Figure 9: Network meta-analysis of first-line eradication of *H pylori* – relative effect of all options compared with placebo

Table 18: Network meta-analysis of first-line eradication of *H pylori* – rankings for each comparator

Regimen	Including regimens with unlicensed components		Excluding regimens with unlicensed components	
	Probability best	Median rank (95%CrI)	Probability best	Median rank (95%CrI)
BIS-NIT-PPI-TET	0.032	4 (1, 11)	0.152	3 (1, 9)
AMO-PPI-QUI	0.163	4 (1, 14)	0.309	3 (1, 12)
BIS-CLA-NIT	0.133	4 (1, 13)	0.267	3 (1, 12)
AMO-CLA-PPI	0.001	6 (3, 10)	0.024	4 (2, 8)
AMO-NIT-PPI	0.005	7 (2, 12)	0.021	5 (2, 10)
BIS-NTF-TET	0.036	8 (1, 14)	0.016	6 (2, 10)
AMO-H_2RA-NIT	0.043	8 (1, 14)	0.103	7 (1, 12)
BIS-H_2RA-NIT-TET	0.017	8 (2, 14)	0.053	7 (1, 12)
BIS-H_2RA-CLA	0.013	10 (2, 14)	0.042	8 (1, 12)
BIS-NIT-TET	0.002	10 (3, 14)	0.010	8 (2, 12)
CLA-PPI	0.001	13 (6, 14)	0.002	11 (5, 12)
AMO-PPI	0.000	13 (6, 14)	0.002	11 (5, 12)
PPI	0.000	15 (15, 15)	0.000	13 (13, 13)
NTZ-PPI-QUI-TET	0.550	1 (1, 8)		
CLA-NIT-PPI	0.003	8 (3, 12)		

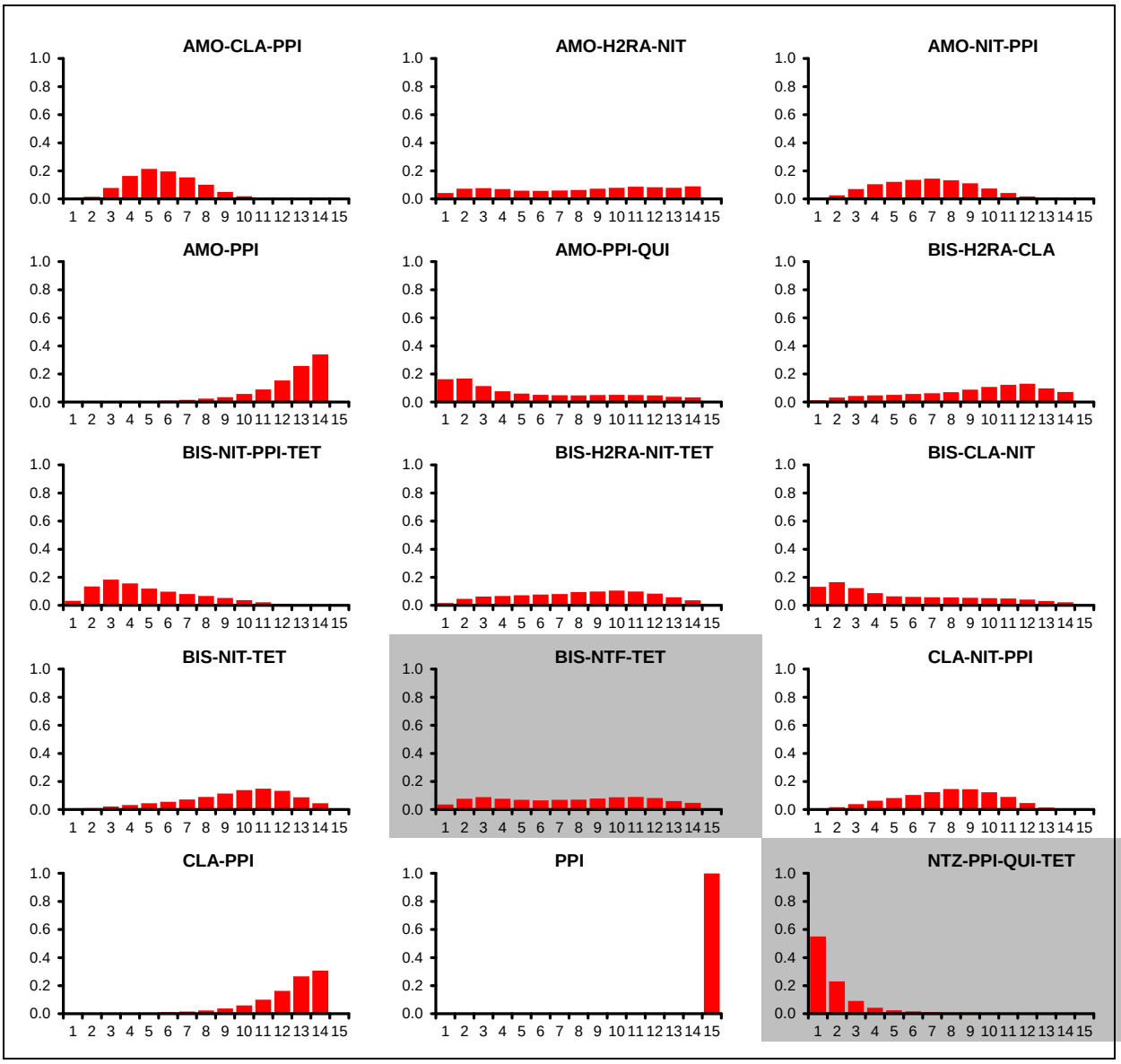


Figure 10: Network meta-analysis of first-line eradication of *H pylori* – rank probability histograms (including regimens with an unlicensed component)

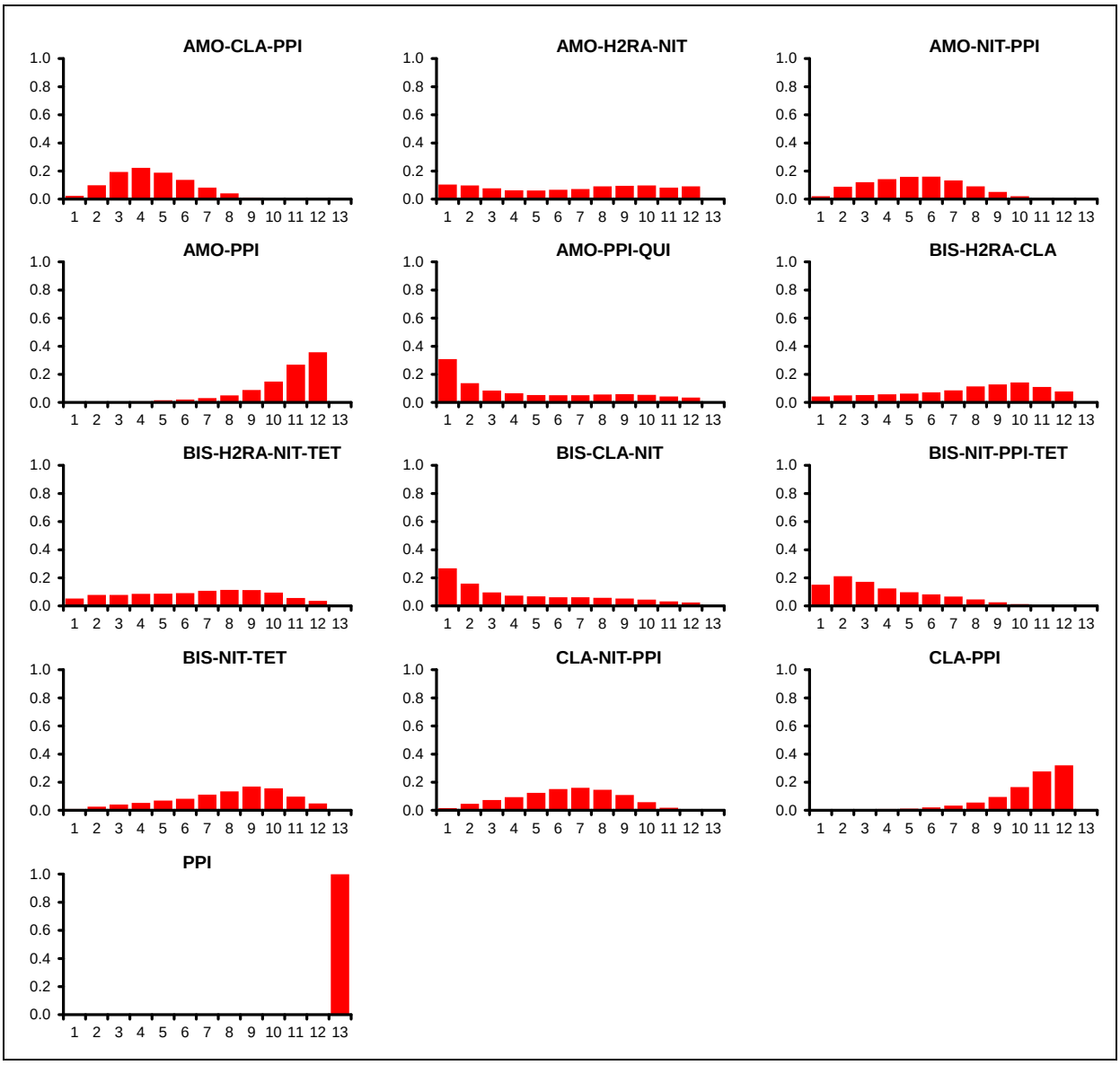


Figure 11: Network meta-analysis of first-line eradication of *H pylori* – rank probability histograms (excluding regimens with an unlicensed component)

Table 19: Network meta-analysis of first-line eradication of *H pylori* – model fit statistics

Residual deviance	Dbar	Dhat	pD	DIC	tau
44.7 (compared to 41 datapoints)	209.322	172.411	36.911	246.234	0.627 (95%CrI: 0.224, 1.406)

E.4.2.2 Second-line eradication

Eradication

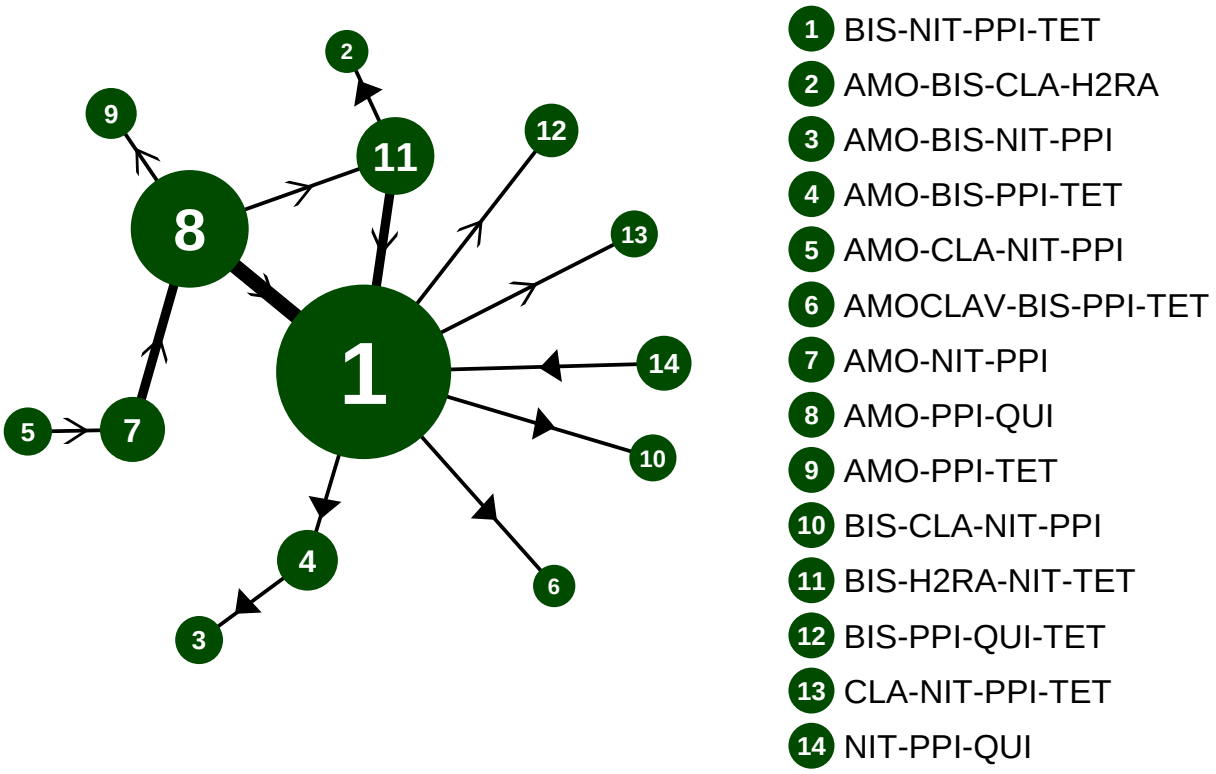
A total of 18 RCTs from the evidence review met the inclusion criteria for the eradication network. Four studies which reported eradication could not be included in the NMA for the following reason:

- The studies compared the same regimens and only the dose or duration differed (internal loop: denoted by a dashed line on the full network diagram [see main guideline])

The data from these studies was analysed using pairwise meta-analyses and was presented to the GDG and considered alongside the outputs from the NMA. Full GRADE tables can be viewed in appendix F

Table 21 is the odds ratio matrix for second-line *H pylori* eradication summarising the results of the conventional pairwise meta-analyses together with the results generated by the NMA for every possible treatment comparison.

The plot of relative effectiveness (Figure 13) shows indirect estimates of interventions in comparison to PPI/BIS/NIT/TET from the NMA with 95% CrI (solid error bars) and direct pairwise head-to-head comparisons with 95% CI (dashed error bars) in graphical form. The rankograms (Figure 14) show the probability of being the best regimen for achieving *H pylori* eradication second-line.



Size of nodes is proportional to total number of participants randomised to receive the treatment in question across the evidence-base. Width of connecting lines is proportional to number of trial-level comparisons available. Arrowheads indicate direction of effect in pairwise data (a > b denotes a is more effective than b) – filled arrowheads show comparisons where one option is significantly superior (p<0.05); outlined arrowheads show direction of trend where effect does not reach statistical significance.

Figure 12: Network meta-analysis of second-line eradication of *H pylori* – evidence network

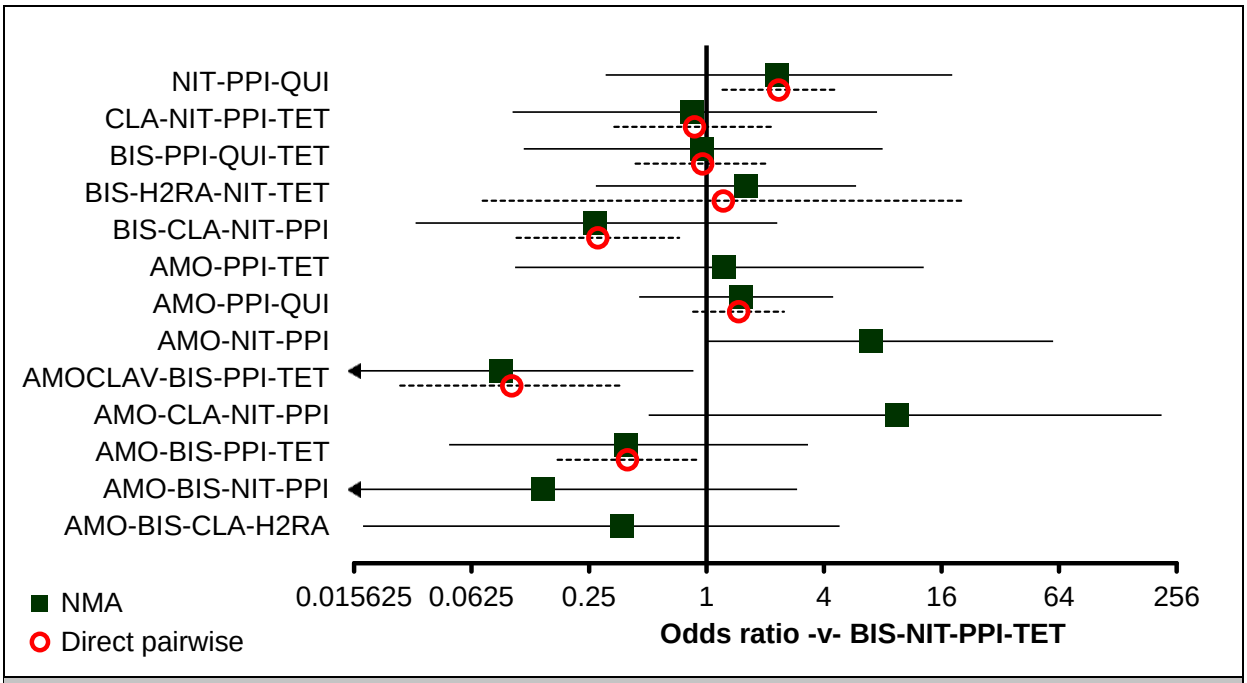
Table 20: Network meta-analysis of second-line eradication of *H pylori* – input data

	BIS-NIT-PPI-TET	AMO-BIS-CLA-H_2RA	AMO-BIS-NIT-PPI	AMO-BIS-PPI-TET	AMO-CLA-NIT-PPI	AMOCLAV-BIS-PPI-TET	AMO-NIT-PPI	AMO-PPI-QUI	AMO-PPI-TET	BIS-CLA-NIT-PPI	BIS-H_2RA-NIT-TET	BIS-PPI-QUI-TET	CLA-NIT-PPI-TET	NIT-PPI-QUI
Bago J et al. (2009)	42/78													60/82
Cheon JH et al. (2006)	19/29					4/25								
Cheon JH et al. (2006)	24/44							31/41						
Chi CH et al. (2003)			29/50	39/50										
Chuah SK et al. (2012)								50/64	48/64					
Chuah S-K et al. (2012)	59/74											60/76		
Georgopoulos SD et al. (2002)	41/49									27/46				
Gisbert JP et al. (1999)	17/30										25/30			
Gisbert JP et al. (2007)								34/50			34/50			
Hu TH et al. (2011)							38/45	31/45						
Koksal AS et al. (2005)		17/28									24/28			
Kuo CH et al. (2009)	53/83							58/83						
Kuo C-H et al. (2013)	43/50							43/51						
Matsumoto Y et al. (2005)							29/30	21/30						
Michopoulos S et al. (2000)	76/76										74/76			
Ueki N et al. (2009)					45/52		43/52							
Wu DC et al. (2006)	36/47												34/46	
Wu DC et al. (2011)	50/62			36/58										

Table 21: Network meta-analysis of second-line eradication of *H pylori* – relative effectiveness of all pairwise combinations

	BIS-NIT-PPI-TET	AMO-BIS-CLA-H_2RA	AMO-BIS-NIT-PPI	AMO-BIS-PPI-TET	AMO-CLA-NIT-PPI	AMOCLAV-BIS-PPI-TET	AMO-NIT-PPI	AMO-PPI-QUI	AMO-PPI-TET	BIS-CLA-NIT-PPI	BIS-H_2RA-NIT-TET	BIS-PPI-QUI-TET	CLA-NIT-PPI-TET	NIT-PPI-QUI
BIS-NIT-PPI-TET		-	-	0.39 (0.17, 0.89)	-	0.10 (0.03, 0.37)	-	1.46 (0.85, 2.50)	-	0.28 (0.11, 0.72)	1.22 (0.07, 20.81)	0.95 (0.43, 2.10)	0.87 (0.34, 2.22)	2.34 (1.21, 4.53)
AMO-BIS-CLA-H_2RA	0.37 (0.02, 4.81)		-	-	-	-	-	-	-	-	3.88 (1.06, 14.28)	-	-	-
AMO-BIS-NIT-PPI	0.15 (0.01, 2.92)	0.39 (0.01, 30.56)		2.57 (1.07, 6.15)	-	-	-	-	-	-	-	-	-	-
AMO-BIS-PPI-TET	0.39 (0.05, 3.31)	1.04 (0.04, 45.65)	2.63 (0.31, 21.35)		-	-	-	-	-	-	-	-	-	-
AMO-CLA-NIT-PPI	9.46 (0.50, 214.6)	26.51 (0.67, 1907)	63.92 (0.98, 4838)	24.36 (0.66, 1142)		-	0.74 (0.25, 2.17)	-	-	-	-	-	-	-
AMOCLAV-BIS-PPI-TET	0.09 (0.01, 0.86)	0.24 (0.01, 11.19)	0.60 (0.01, 25.49)	0.23 (0.01, 4.85)	0.01 (0.00, 0.40)		-	-	-	-	-	-	-	-
AMO-NIT-PPI	6.96 (0.99, 59.61)	18.88 (0.95, 720.1)	46.53 (1.36, 1980)	17.95 (1.06, 388.4)	0.73 (0.08, 6.83)	79.82 (3.98, 1997)		0.24 (0.05, 1.07)	-	-	-	-	-	-
AMO-PPI-QUI	1.51 (0.45, 4.45)	4.05 (0.28, 83.03)	10.18 (0.36, 225.5)	3.89 (0.31, 38.64)	0.16 (0.01, 2.22)	16.73 (1.23, 223.5)	0.22 (0.03, 1.01)		0.84 (0.37, 1.91)	-	1.00 (0.43, 2.32)	-	-	-
AMO-PPI-TET	1.23 (0.10, 12.97)	3.28 (0.11, 137.2)	8.59 (0.17, 353.9)	3.20 (0.12, 72.66)	0.13 (0.00, 3.88)	13.87 (0.48, 425.4)	0.18 (0.01, 2.28)	0.82 (0.10, 6.88)		-	-	-	-	-
BIS-CLA-NIT-PPI	0.27 (0.03, 2.31)	0.73 (0.03, 32.33)	1.81 (0.04, 70.96)	0.69 (0.03, 13.82)	0.03 (0.00, 1.09)	3.04 (0.13, 73.85)	0.04 (0.00, 0.71)	0.18 (0.02, 2.19)	0.22 (0.01, 6.33)		-	-	-	-
BIS-H_2RA-NIT-TET	1.60 (0.27, 5.83)	4.31 (0.42, 48.07)	10.95 (0.27, 231.0)	4.15 (0.23, 41.50)	0.17 (0.00, 3.14)	17.93 (0.89, 241.1)	0.23 (0.01, 1.72)	1.07 (0.18, 4.29)	1.30 (0.07, 15.30)	6.01 (0.32, 63.89)		-	-	-
BIS-PPI-QUI-TET	0.95 (0.12, 7.97)	2.57 (0.10, 103.8)	6.56 (0.16, 240.7)	2.48 (0.12, 46.05)	0.10 (0.00, 3.79)	10.61 (0.50, 264.4)	0.14 (0.01, 2.25)	0.63 (0.06, 7.47)	0.77 (0.03, 20.53)	3.48 (0.18, 73.54)	0.58 (0.06, 10.78)		-	-
CLA-NIT-PPI-TET	0.84 (0.10, 7.48)	2.29 (0.09, 99.93)	5.96 (0.14, 224.4)	2.22 (0.11, 44.82)	0.09 (0.00, 3.35)	9.73 (0.42, 236.9)	0.12 (0.01, 2.10)	0.57 (0.05, 7.07)	0.70 (0.03, 18.60)	3.20 (0.16, 65.93)	0.53 (0.05, 9.65)	0.90 (0.05, 18.63)		-
NIT-PPI-QUI	2.31 (0.30, 18.15)	6.14 (0.25, 256.9)	16.17 (0.41, 557.4)	6.07 (0.29, 116.9)	0.25 (0.01, 8.62)	25.88 (1.28, 624.8)	0.34 (0.02, 5.43)	1.54 (0.16, 16.77)	1.90 (0.08, 49.36)	8.77 (0.44, 176.3)	1.43 (0.14, 23.78)	2.44 (0.12, 45.40)	2.72 (0.14, 51.83)	

Values given are odds ratios.
The segment below and to the left of the shaded diagonal is derived from the network meta-analysis, reflecting direct and indirect evidence of treatment effects (row versus column). The point estimate reflects the mean of the posterior distribution, and numbers in parentheses are 95% credible intervals.
The segment above and to the right of the shaded diagonal gives pooled direct evidence (random-effects pairwise meta-analysis), where available (column versus row). Numbers in parentheses are 95% confidence intervals.



Values less than 1 favour BIS-NIT-PPI-TET; values greater than 1 favour the comparator treatment. Solid error bars are 95% credible intervals while dashed error bars are 95% confidence interval.

Figure 13: Network meta-analysis of second-line eradication of *H pylori* – relative effect of all options compared with BIS-NIT-PPI-TET

Table 22: Network meta-analysis of second-line eradication of *H pylori* – rankings for each comparator

	Probability best	Median rank (95%CrI)
AMO-CLA-NIT-PPI	0.569	1 (1, 9)
AMO-NIT-PPI	0.261	2 (1, 6)
NIT-PPI-QUI	0.078	4 (1, 11)
BIS-H_2RA-NIT-TET	0.012	5 (2, 11)
AMO-PPI-QUI	0.001	5 (3, 10)
AMO-PPI-TET	0.024	6 (2, 13)
BIS-NIT-PPI-TET	0.000	7 (4, 10)
BIS-PPI-QUI-TET	0.019	8 (2, 13)
CLA-NIT-PPI-TET	0.018	8 (2, 13)
AMO-BIS-PPI-TET	0.004	11 (3, 13)
AMO-BIS-CLA-H_2RA	0.007	11 (3, 14)
BIS-CLA-NIT-PPI	0.002	11 (4, 14)
AMO-BIS-NIT-PPI	0.005	13 (4, 14)
AMOCLAV-BIS-PPI-TET	0.000	13 (8, 14)

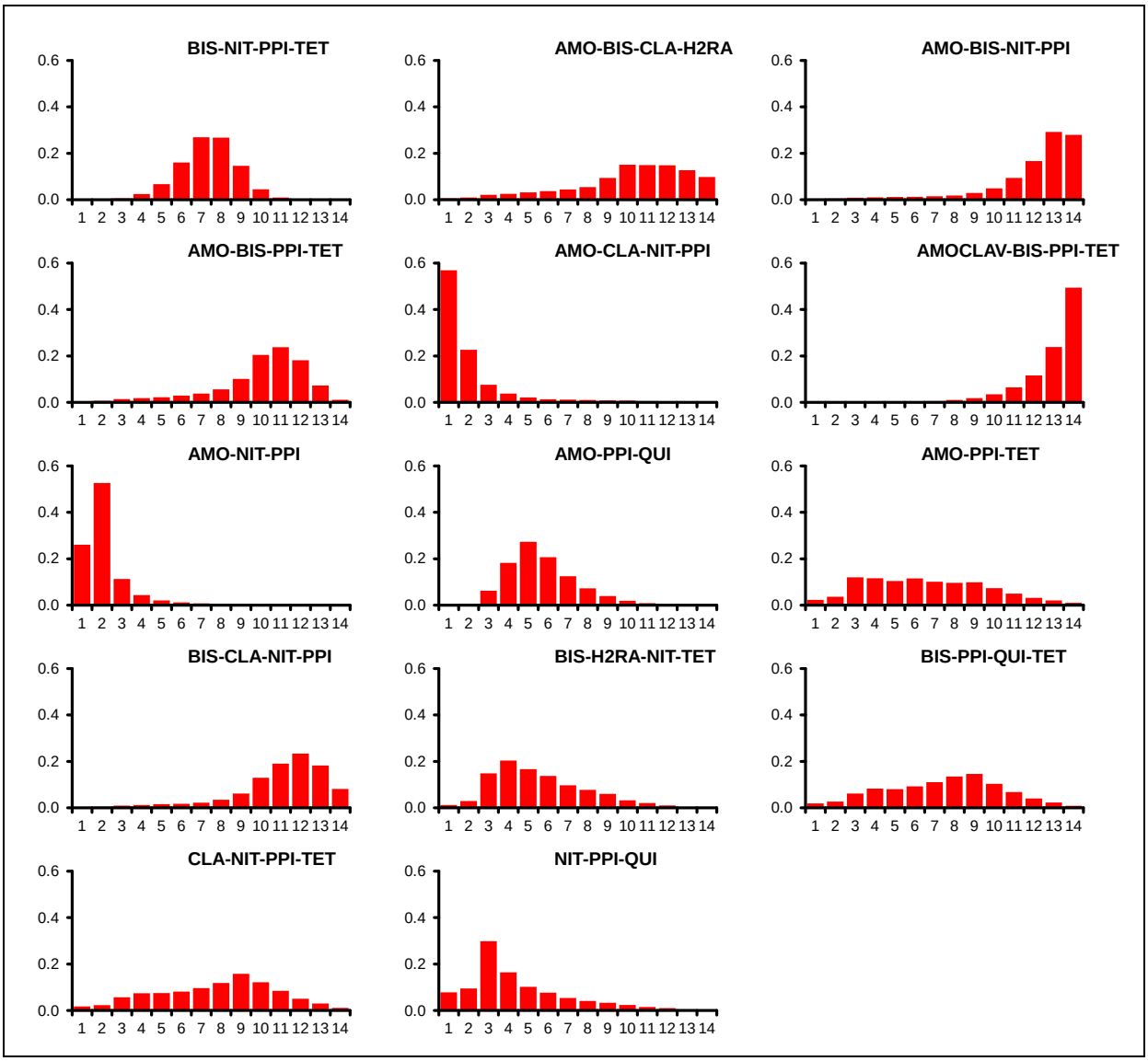


Figure 14: Network meta-analysis of second-line eradication of *H pylori* – rank probability histograms

Table 23: Network meta-analysis of second-line eradication of *H pylori* – model fit statistics

Residual deviance	Dbar	Dhat	pD	DIC	tau
38.76 (compared to 36 datapoints)	176.376	142.125	34.251	210.626	0.678 (95%CrI: 0.045, 1.854)

Second-line adverse events – rash

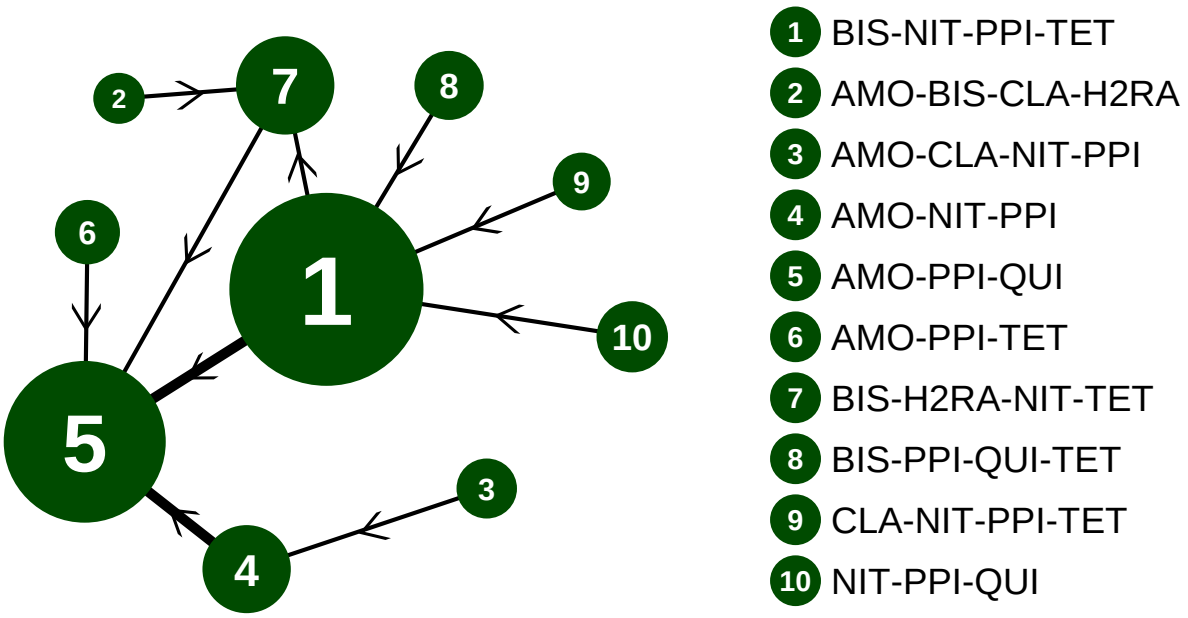
A total of 12 RCTs from the evidence review met the inclusion criteria for the adverse events (rash) network. One study which reported on rash could not be included in the NMA for the following reason:

- The study compared the same regimens and only the duration differed (internal loop: denoted by a dashed line on the full network diagram [see full guideline])

The data from this study were analysed using pairwise meta-analysis and were presented to the GDG and considered alongside the outputs from the NMA. Full GRADE tables can be viewed in appendix F.

Table 25 is the odds ratio matrix for rash summarising the results of the conventional pairwise meta-analyses together with the results generated by the NMA for every possible treatment comparison.

The plot of relative effectiveness (Figure 16) shows indirect estimates of interventions in comparison to (PPI/BIS/NIT/TET) from the NMA with 95% CrI (solid error bars) and direct pairwise head-to-head comparisons with 95% CIs (dashed error bars) in graphical form. The rankograms (Figure 17) show the probability of being the best second-line eradication regimen for achieving the lowest incidence of rash.



Size of nodes is proportional to total number of participants randomised to receive the treatment in question across the evidence-base. Width of connecting lines is proportional to number of trial-level comparisons available. Arrowheads indicate direction of effect in pairwise data (a > b denotes a is more effective than b) – filled arrowheads show comparisons where one option is significantly superior (p<0.05); outlined arrowheads show direction of trend where effect does not reach statistical significance.

Figure 15: Network meta-analysis of second-line *H pylori* treatment – rash – evidence network

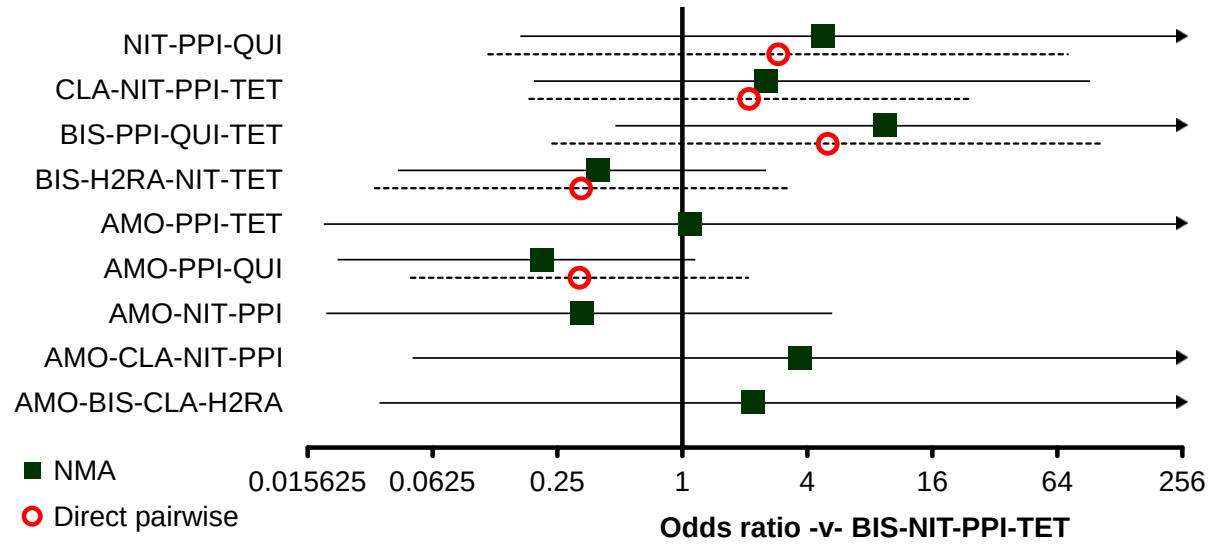
Table 24: Network meta-analysis of second-line *H pylori* treatment – rash – input data

	BIS-NIT-PPI-TET	AMO-BIS-CLA-H_2RA	AMO-CLA-NIT-PPI	AMO-NIT-PPI	AMO-PPI-QUI	AMO-PPI-TET	BIS-H_2RA-NIT-TET	BIS-PPI-QUI-TET	CLA-NIT-PPI-TET	NIT-PPI-QUI
Bago,J. et al. (2009)	0/78									1/82
Chuah,S.K. et al. (2012)					0/64	1/64				
Chuah,S.-K. et al. (2012)	0/74							2/76		
Gisbert,J.P. et al. (2007)					0/50		1/50			
Hu,T.H. et al. (2011)				2/45	0/45					
Koksal,A.S. et al. (2005)		1/28					0/28			
Kuo,C.H. et al. (2009)	1/83				0/83					
Kuo,C.-H. et al. (2013)	3/50				1/51					
Matsumoto,Y. et al. (2005)				0/30	1/30					
Michopoulos,S. et al. (2000)	3/76						1/76			
Ueki,N. et al. (2009)			2/52	0/52						
Wu,D.C. et al. (2006)	1/47								2/46	

Table 25: Network meta-analysis of second-line *H pylori* treatment – rash – relative effectiveness of all pairwise combinations

	BIS-NIT-PPI-TET	AMO-BIS-CLA-H_2RA	AMO-CLA-NIT-PPI	AMO-NIT-PPI	AMO-PPI-QUI	AMO-PPI-TET	BIS-H_2RA-NIT-TET	BIS-PPI-QUI-TET	CLA-NIT-PPI-TET	NIT-PPI-QUI
BIS-NIT-PPI-TET		-	-	-	0.32 (0.05, 2.07)	-	0.32 (0.03, 3.19)	5.00 (0.24, 105.93)	2.09 (0.18, 23.89)	2.89 (0.12, 72.00)
AMO-BIS-CLA-H_2RA	2.19 (0.03, 1444)		-	-	-	-	0.32 (0.01, 8.24)	-	-	-
AMO-CLA-NIT-PPI	3.68 (0.05, 2526)	1.70 (0.00, 3041)		0.19 (0.01, 4.11)	-	-	-	-	-	-
AMO-NIT-PPI	0.33 (0.02, 5.26)	0.15 (0.00, 16.53)	0.10 (0.00, 2.06)		0.67 (0.11, 4.05)	-	-	-	-	-
AMO-PPI-QUI	0.21 (0.02, 1.15)	0.09 (0.00, 6.27)	0.06 (0.00, 2.30)	0.62 (0.07, 4.00)		3.05 (0.12, 76.21)	3.06 (0.12, 76.95)	-	-	-
AMO-PPI-TET	1.09 (0.02, 555.3)	0.51 (0.00, 543.70)	0.31 (0.00, 297.90)	3.39 (0.06, 1668)	5.17 (0.17, 2110)		-	-	-	-
BIS-H_2RA-NIT-TET	0.39 (0.04, 2.53)	0.19 (0.00, 5.93)	0.10 (0.00, 9.01)	1.18 (0.05, 23.09)	1.90 (0.20, 20.81)	0.34 (0.00, 23.39)		-	-	-
BIS-PPI-QUI-TET	9.48 (0.48, 3450)	4.56 (0.00, 5974)	2.83 (0.00, 3169)	32.29 (0.45, 16,050)	50.34 (1.37, 21,960)	9.36 (0.01, 10170)	26.35 (0.71, 12,950)		-	-
CLA-NIT-PPI-TET	2.54 (0.19, 92.10)	1.17 (0.00, 261.6)	0.71 (0.00, 169.80)	8.08 (0.18, 711.50)	12.89 (0.55, 863.6)	2.31 (0.00, 549.70)	6.86 (0.26, 435.40)	0.25 (0.00, 25.86)		-
NIT-PPI-QUI	4.78 (0.17, 2989)	2.34 (0.00, 4092)	1.41 (0.00, 2623)	16.23 (0.17, 16060)	25.45 (0.51, 19,200)	4.82 (0.00, 7237)	13.19 (0.26, 10,980)	0.49 (0.00, 656.70)	1.97 (0.01, 1835)	

Values given are odds ratios.
The segment below and to the left of the shaded diagonal is derived from the network meta-analysis, reflecting direct and indirect evidence of treatment effects (row versus column). The point estimate reflects the mean of the posterior distribution, and numbers in parentheses are 95% credible intervals.
The segment above and to the right of the shaded diagonal gives pooled direct evidence (random-effects pairwise meta-analysis), where available (column versus row). Numbers in parentheses are 95% confidence intervals.



Values greater than 1 favour BIS-NIT-PPI-TET; values less than 1 favour the comparator treatment. Solid error bars are 95% credible intervals while dashed error bars are 95% confidence interval.

Figure 16: Network meta-analysis of second-line *H pylori* treatment – rash – relative effect of all options compared with BIS-NIT-PPI-TET

Table 26: Network meta-analysis of second-line *H pylori* treatment – rash – rankings for each comparator

	Probability best	Median rank (95%CrI)
AMO-PPI-QUI	0.339	2 (1, 5)
BIS-H_2RA-NIT-TET	0.161	3 (1, 7)
AMO-NIT-PPI	0.196	3 (1, 8)
AMO-PPI-TET	0.126	5 (1, 10)
BIS-NIT-PPI-TET	0.006	5 (2, 8)
AMO-BIS-CLA-H_2RA	0.081	7 (1, 10)
CLA-NIT-PPI-TET	0.027	7 (1, 10)
AMO-CLA-NIT-PPI	0.031	8 (1, 10)
NIT-PPI-QUI	0.029	8 (1, 10)
BIS-PPI-QUI-TET	0.006	9 (3, 10)

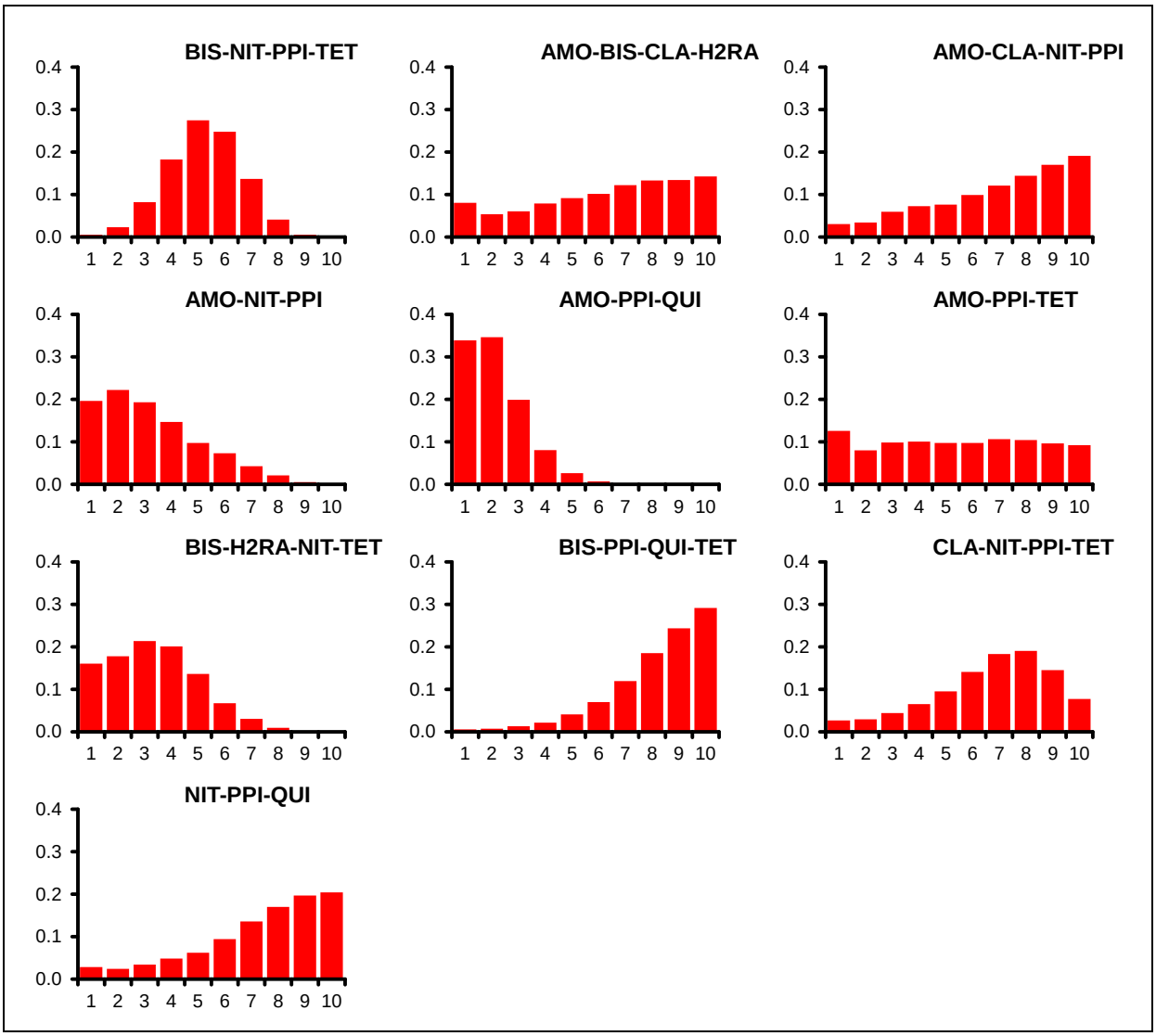


Figure 17: Network meta-analysis of second-line *H pylori* treatment – rash – rank probability histograms

Table 27: Network meta-analysis of second-line *H pylori* treatment – rash – model fit statistics

Residual deviance	Dbar	Dhat	pD	DIC	tau
25.72 (compared to 24 datapoints)	75.346	56.733	18.613	93.959	n/a (fixed-effects model)

Second-line adverse events – loose stools

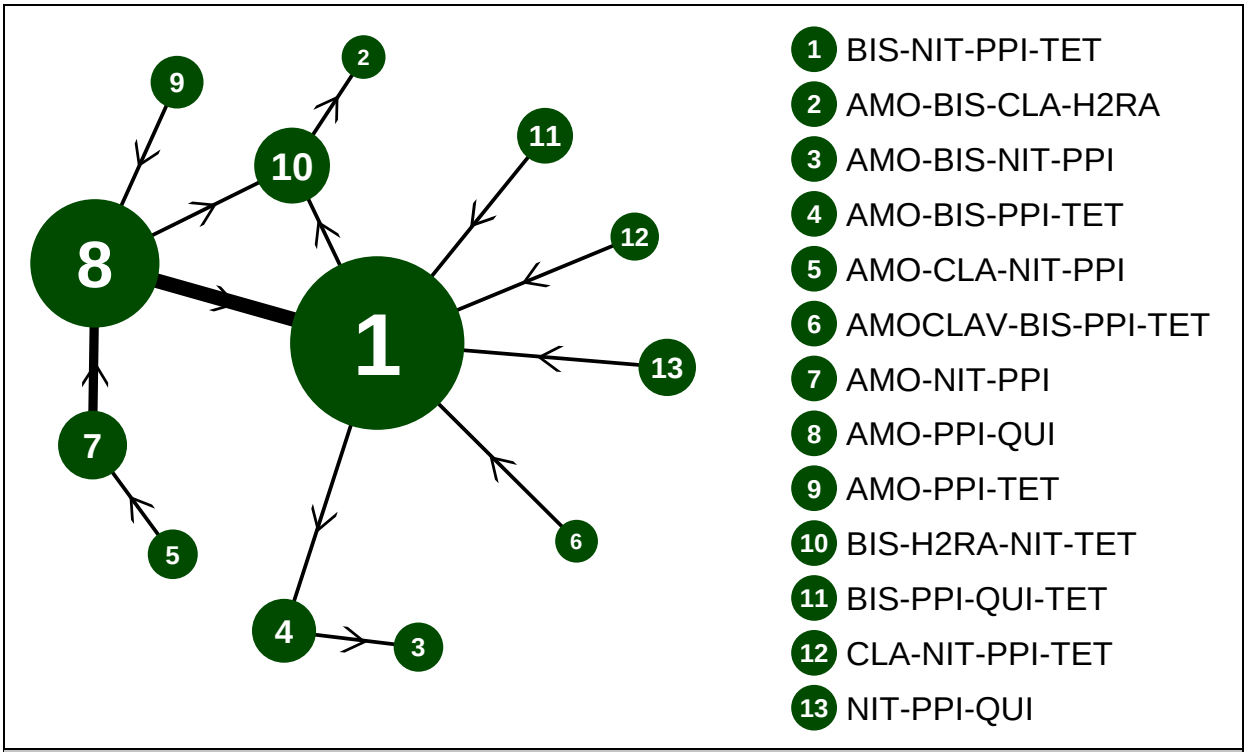
A total of 16 RCTs from the evidence review met the inclusion criteria for the adverse events (loose stools) network. Three studies which reported on loose stools could not be included in the NMA for the following reason:

- The studies compared the same regimens and only the dose or duration differed (internal loop: denoted by a dashed line on the full network diagram [see full guideline])

The data from these studies was analysed using pairwise meta-analyses and was presented to the GDG and considered alongside the outputs from the NMA. Full GRADE tables can be viewed in appendix F

Table 29 is the odds ratio matrix for loose stools summarising the results of the conventional pairwise meta-analyses together with the results generated by the NMA for every possible treatment comparison.

The plot of relative effectiveness (Figure 19) shows indirect estimates of interventions in comparison to (PPI/BIS/NIT/TET) and direct pairwise head-to-head comparisons with 95% CI (dashed error bars) in graphical form. The rankograms (Figure 20) show the probability of being the best second-line eradication regimen for achieving the lowest incidence of loose stools.



Size of nodes is proportional to total number of participants randomised to receive the treatment in question across the evidence-base. Width of connecting lines is proportional to number of trial-level comparisons available. Arrowheads indicate direction of effect in pairwise data (a > b denotes a is more effective than b) – filled arrowheads show comparisons where one option is significantly superior (p<0.05); outlined arrowheads show direction of trend where effect does not reach statistical significance.

Figure 18: Network meta-analysis of second-line *H pylori* treatment – loose stools – evidence network

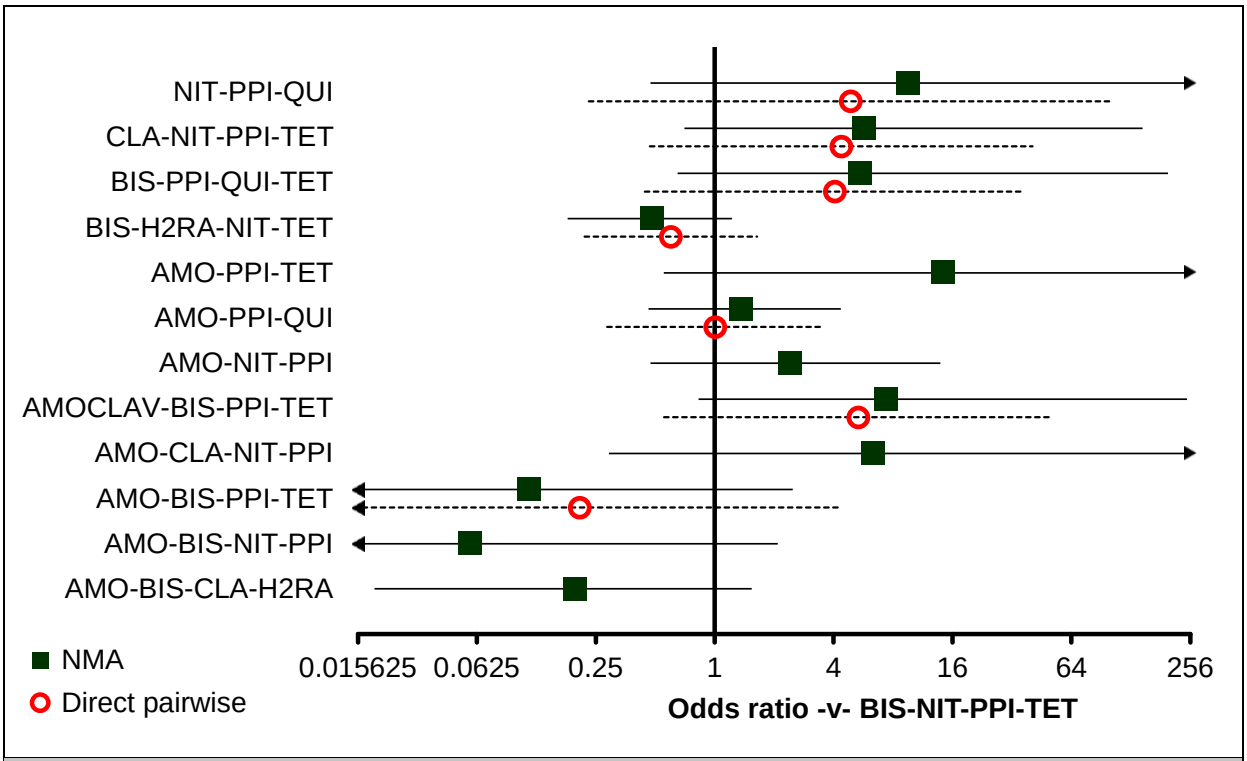
Table 28: Network meta-analysis of second-line *H pylori* treatment – loose stools – input data

	BIS-NIT-PPI-TET	AMO-BIS-CLA-H_2RA	AMO-BIS-NIT-PPI	AMO-BIS-PPI-TET	AMO-CLA-NIT-PPI	AMOCLAV-BIS-PPI-TET	AMO-NIT-PPI	AMO-PPI-QUI	AMO-PPI-TET	BIS-H_2RA-NIT-TET	BIS-PPI-QUI-TET	CLA-NIT-PPI-TET	NIT-PPI-QUI
Bago,J. et al. (2009)	0/78												2/82
Cheon,J.H. et al. (2006)	1/29					4/25							
Cheon,J.H. et al. (2006)	0/44							1/41					
Chi,C.H. et al. (2003)			3/50	5/50									
Chuah,S.K. et al. (2012)								0/64	2/64				
Chuah,S.-K. et al. (2012)	1/74										4/76		
Gisbert,J.P. et al. (2007)								5/50		1/50			
Hu,T.H. et al. (2011)							2/45	2/45					
Koksal,A.S. et al. (2005)		2/28								4/28			
Kuo,C.H. et al. (2009)	2/83							0/83					
Kuo,C.-H. et al. (2013)	2/50							3/51					
Matsumoto,Y. et al. (2005)							6/30	3/30					
Michopoulos,S. et al. (2000)	11/76									7/76			
Ueki,N. et al. (2009)					2/52		1/52						
Wu,D.C. et al. (2006)	1/47											4/46	
Wu,D.C. et al. (2011)	2/62			0/58									

Table 29: Network meta-analysis of second-line *H pylori* treatment – loose stools – relative effectiveness of all pairwise combinations

	BIS-NIT-PPI-TET	AMO-BIS-CLA-H_2RA	AMO-BIS-NIT-PPI	AMO-BIS-PPI-TET	AMO-CLA-NIT-PPI	AMOCLAV-BIS-PPI-TET	AMO-NIT-PPI	AMO-PPI-QUI	AMO-PPI-TET	BIS-H_2RA-NIT-TET	BIS-PPI-QUI-TET	CLA-NIT-PPI-TET	NIT-PPI-QUI
BIS-NIT-PPI-TET		-	-	0.21 (0.01, 4.40)	-	5.33 (0.55, 51.27)	-	1.00 (0.29, 3.53)	-	0.60 (0.22, 1.64)	4.06 (0.44, 37.17)	4.38 (0.47, 40.78)	4.88 (0.23, 103.1)
AMO-BIS-CLA-H2RA	0.20 (0.02, 1.54)		-	-	-	-	-	-	-	2.17 (0.36, 12.92)	-	-	-
AMO-BIS-NIT-PPI	0.06 (0.00, 2.09)	0.28 (0.00, 23.71)		1.74 (0.39, 7.71)	-	-	-	-	-	-	-	-	-
AMO-BIS-PPI-TET	0.12 (0.00, 2.48)	0.54 (0.00, 30.20)	1.82 (0.40, 10.03)		-	-	-	-	-	-	-	-	-
AMO-CLA-NIT-PPI	6.32 (0.29, 284.0)	33.54 (0.83, 2470)	125.3 (0.96,189400)	66.40 (0.70, 84290)		-	0.49 (0.04, 5.58)	-	-	-	-	-	-
AMOCLAV-BIS-PPI-TET	7.38 (0.83, 246.2)	41.11 (1.76, 2377)	151.2 (1.87,223200)	77.62 (1.35, 94510)	1.24 (0.02, 119.6)		-	-	-	-	-	-	-
AMO-NIT-PPI	2.43 (0.47, 13.90)	12.71 (1.04, 209.9)	44.17 (0.87, 32260)	22.88 (0.67, 14000)	0.40 (0.01, 5.70)	0.32 (0.01, 5.52)		0.59 (0.18, 1.92)	-	-	-	-	-
AMO-PPI-QUI	1.37 (0.46, 4.36)	7.07 (0.78, 87.77)	23.94 (0.58, 16360)	12.43 (0.47, 6976)	0.22 (0.01, 4.03)	0.18 (0.00, 2.28)	0.57 (0.16, 1.88)		5.16 (0.24, 109.6)	0.18 (0.02, 1.63)	-	-	-
AMO-PPI-TET	14.36 (0.55, 5632)	79.68 (1.63, 37600)	311.6 (1.95,1821000)	163.7 (1.37,856700)	2.38 (0.02, 1565)	1.87 (0.02, 1094)	5.65 (0.21, 2399)	9.86 (0.51, 3936)		-	-	-	-
BIS-H2RA-NIT-TET	0.48 (0.18, 1.22)	2.41 (0.40, 21.11)	8.27 (0.21, 5780)	4.26 (0.17, 2328)	0.08 (0.00, 1.80)	0.06 (0.00, 0.73)	0.20 (0.03, 1.09)	0.35 (0.09, 1.18)	0.03 (0.00, 0.91)		-	-	-
BIS-PPI-QUI-TET	5.46 (0.65, 197.7)	30.33 (1.34, 1918)	112.4 (1.39, 151500)	60.23 (1.04, 65290)	0.92 (0.01, 95.92)	0.76 (0.01, 47.80)	2.35 (0.14, 110.4)	4.07 (0.34, 168.4)	0.39 (0.00, 47.87)	11.62 (1.08, 472.6)		-	-
CLA-NIT-PPI-TET	5.72 (0.70, 147.0)	31.98 (1.52, 1437)	114.4 (1.54, 128200)	60.49 (1.17, 55780)	0.96 (0.01, 75.10)	0.77 (0.01, 35.92)	2.41 (0.16, 85.44)	4.20 (0.37, 123.1)	0.40 (0.00, 38.87)	11.99 (1.17, 347.4)	1.04 (0.02, 45.54)		-
NIT-PPI-QUI	9.55 (0.47, 7851)	55.03 (1.24, 58440)	240.8 (1.32, 1514000)	124.8 (0.95, 731700)	1.74 (0.01, 2215)	1.32 (0.01, 1528)	4.12 (0.12, 3877)	7.17 (0.27, 5896)	0.72 (0.00, 850.8)	20.16 (0.85, 16870)	1.79 (0.02, 1884)	1.73 (0.02, 1748)	

Values given are odds ratios.
The segment below and to the left of the shaded diagonal is derived from the network meta-analysis, reflecting direct and indirect evidence of treatment effects (row versus column). The point estimate reflects the mean of the posterior distribution, and numbers in parentheses are 95% credible intervals.
The segment above and to the right of the shaded diagonal gives pooled direct evidence (random-effects pairwise meta-analysis), where available (column versus row). Numbers in parentheses are 95% confidence intervals.



Values greater than 1 favour BIS-NIT-PPI-TET; values less than 1 favour the comparator treatment. Solid error bars are 95% credible intervals while dashed error bars are 95% confidence interval.

Figure 19: Network meta-analysis of second-line *H pylori* treatment – loose stools – relative effect of all options compared with placebo

Table 30: Network meta-analysis of second-line *H pylori* treatment – loose stools – rankings for each comparator

	Probability best	Median rank (95%CrI)
AMO-BIS-NIT-PPI	0.581	1 (1, 7)
AMO-BIS-PPI-TET	0.128	2 (1, 8)
AMO-BIS-CLA-H_2RA	0.262	3 (1, 7)
BIS-H_2RA-NIT-TET	0.017	4 (2, 6)
BIS-NIT-PPI-TET	0.001	5 (3, 8)
AMO-PPI-QUI	0.001	6 (4, 9)
AMO-NIT-PPI	0.001	8 (4, 11)
AMO-CLA-NIT-PPI	0.005	10 (3, 13)
BIS-PPI-QUI-TET	0.001	10 (4, 13)
AMOCLAV-BIS-PPI-TET	0.000	10 (5, 13)
CLA-NIT-PPI-TET	0.000	10 (5, 13)
AMO-PPI-TET	0.001	11 (4, 13)
NIT-PPI-QUI	0.002	11 (4, 13)

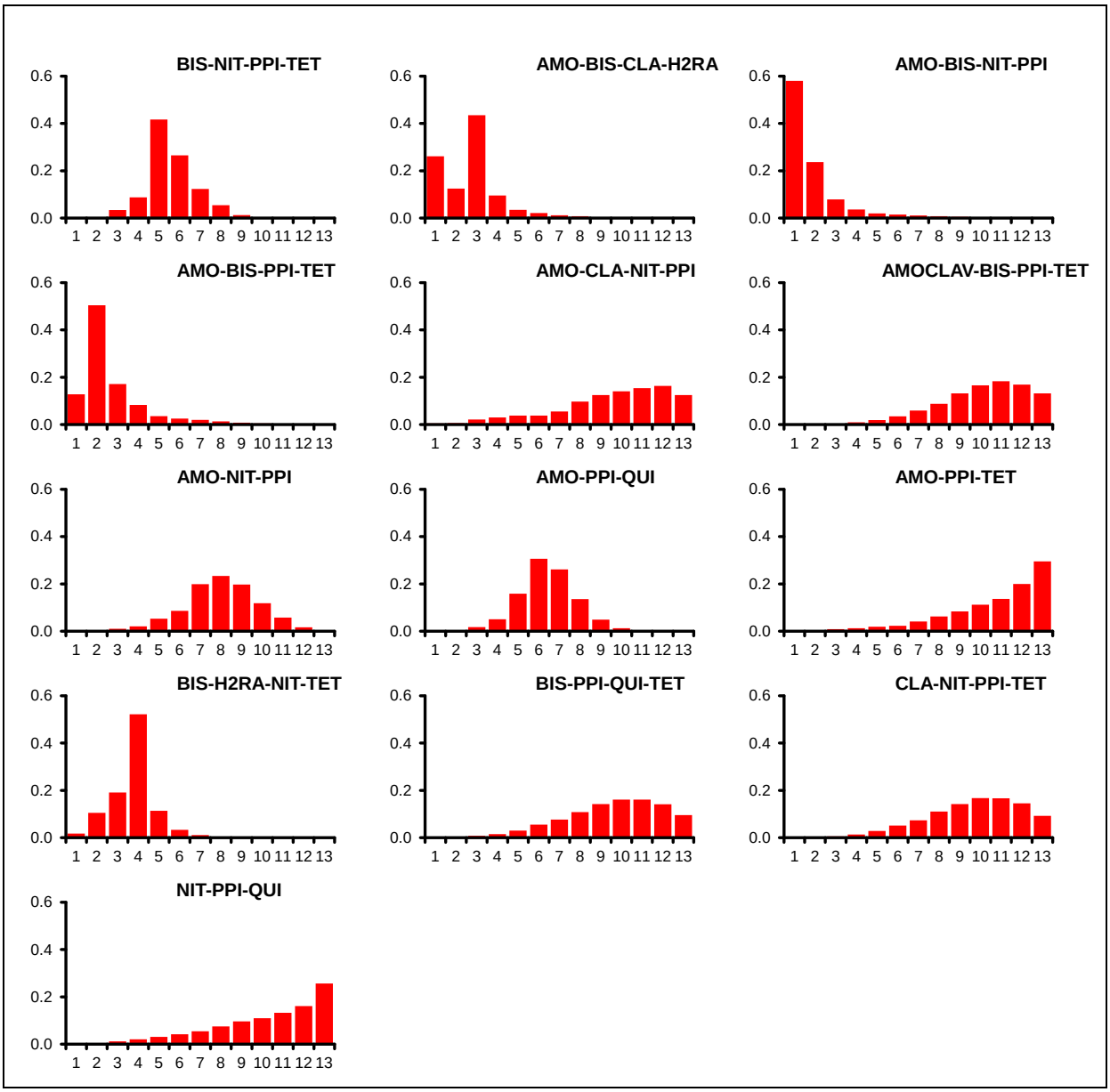


Figure 20: Network meta-analysis of second-line *H pylori* treatment – loose stools – rank probability histograms

Table 31: Network meta-analysis of second-line *H pylori* treatment – loose stools – model fit statistics

Residual deviance	Dbar	Dhat	pD	DIC	tau
34.00 (compared to 32 datapoints)	116.721	90.389	26.332	143.053	n/a (fixed-effects model)

Second-line adherence to medication

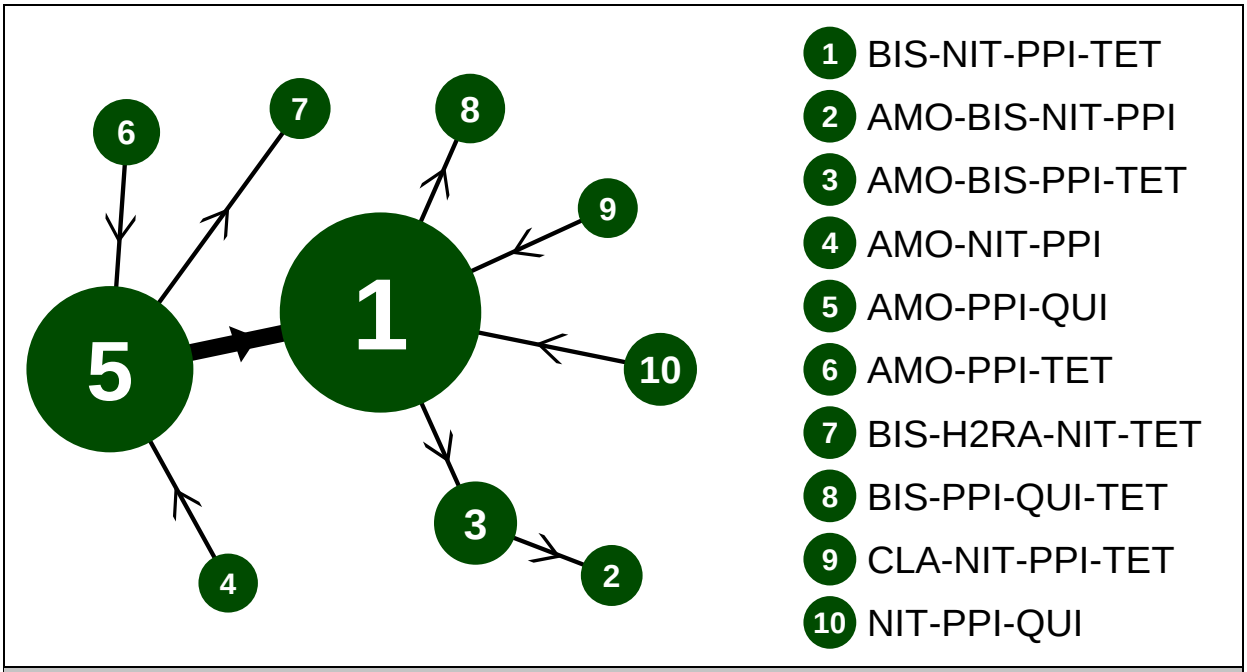
A total of 11 RCTs from the evidence review met the inclusion criteria for the adherence to medication network. Three studies which reported on adherence to medication could not be included in the NMA for the following reason:

- The studies compared the same regimens and only the duration differed (internal loop: denoted by a dashed line on the full network diagram [see full guideline])

The data from this study have been analysed using pairwise meta-analysis and were presented to the GDG and considered alongside the outputs from the NMA. Full GRADE tables can be viewed in appendix F

Table 33 is the odds ratio matrix for adherence to medication summarising the results of the conventional pairwise meta-analyses together with the results generated by the NMA for every possible treatment comparison.

The plot of relative effectiveness (Figure 22) shows indirect estimates of interventions in comparison to (PPI/BIS/NIT/TET) from the NMA with 95% CrI (solid error bars) and direct pairwise head-to-head comparisons with 95% CI (dashed error bars) in graphical form. The rankograms (Figure 23) show the probability of being the regimen with the best adherence.



Size of nodes is proportional to total number of participants randomised to receive the treatment in question across the evidence-base. Width of connecting lines is proportional to number of trial-level comparisons available. Arrowheads indicate direction of effect in pairwise data (a > b denotes a is more effective than b) – filled arrowheads show comparisons where one option is significantly superior (p<0.05); outlined arrowheads show direction of trend where effect does not reach statistical significance.

Figure 21: Network meta-analysis of second-line *H pylori* treatment – adherence to treatment – evidence network

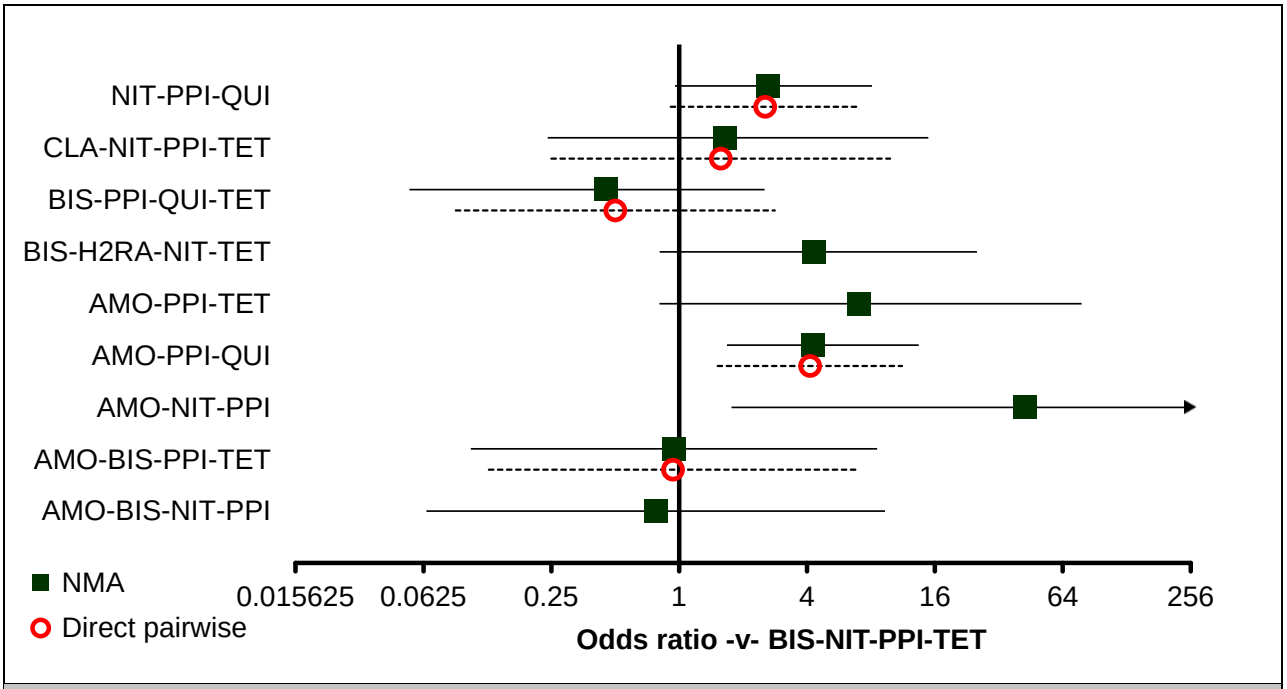
Table 32: Network meta-analysis of second-line *H pylori* treatment – adherence to treatment – input data

	BIS-NIT-PPI-TET	AMO-BIS-NIT-PPI	AMO-BIS-PPI-TET	AMO-NIT-PPI	AMO-PPI-QUI	AMO-PPI-TET	BIS-H_2RA-NIT-TET	BIS-PPI-QUI-TET	CLA-NIT-PPI-TET	NIT-PPI-QUI
Bago,J. et al. (2009)	65/78									76/82
Cheon,J.H. et al. (2006)	33/44				37/41					
Chi,C.H. et al. (2003)		43/50	44/50							
Chuah,S.K. et al. (2012)					61/64	62/64				
Chuah,S.-K. et al. (2012)	69/71							69/73		
Gisbert,J.P. et al. (2007)					45/50		45/50			
Hu,T.H. et al. (2011)				45/45	43/45					
Kuo,C.H. et al. (2009)	66/71				79/80					
Kuo,C.-H. et al. (2013)	47/50				51/51					
Wu,D.C. et al. (2006)	43/46								45/47	
Wu,D.C. et al. (2011)	60/62		56/58							

Table 87: Network meta-analysis of second-line *H pylori* treatment – adherence to treatment – relative effectiveness of all pairwise combinations

	BIS-NIT-PPI-TET	AMO-BIS-NIT-PPI	AMO-BIS-PPI-TET	AMO-NIT-PPI	AMO-PPI-QUI	AMO-PPI-TET	BIS-H_2RA-NIT-TET	BIS-PPI-QUI-TET	CLA-NIT-PPI-TET	NIT-PPI-QUI
BIS-NIT-PPI-TET		-	0.93 (0.13, 6.85)	-	4.13 (1.52, 11.22)	-	-	0.50 (0.09, 2.82)	1.57 (0.25, 9.86)	2.53 (0.91, 7.04)
AMO-BIS-NIT-PPI	0.78 (0.06, 9.31)		1.19 (0.37, 3.84)	-	-	-	-	-	-	-
AMO-BIS-PPI-TET	0.95 (0.10, 8.56)	1.20 (0.36, 4.09)		-	-	-	-	-	-	-
AMO-NIT-PPI	42.69 (1.76, 18480)	59.89 (0.96, 36650)	48.81 (0.94, 28460)		0.19 (0.01, 4.10)	-	-	-	-	-
AMO-PPI-QUI	4.29 (1.68, 13.41)	5.56 (0.38, 84.69)	4.59 (0.41, 55.14)	0.11 (0.00, 2.12)		1.52 (0.25, 9.45)	1.00 (0.27, 3.69)	-	-	-
AMO-PPI-TET	7.07 (0.81, 78.39)	9.31 (0.33, 289.40)	7.65 (0.34, 198.30)	0.16 (0.00, 6.87)	1.61 (0.23, 14.02)		-	-	-	-
BIS-H_2RA-NIT-TET	4.32 (0.81, 25.24)	5.63 (0.27, 118.10)	4.66 (0.29, 78.75)	0.10 (0.00, 3.01)	1.00 (0.25, 3.99)	0.61 (0.05, 6.66)		-	-	-
BIS-PPI-QUI-TET	0.46 (0.05, 2.53)	0.56 (0.02, 12.56)	0.47 (0.02, 8.00)	0.01 (0.00, 0.42)	0.10 (0.01, 0.75)	0.06 (0.00, 0.98)	0.10 (0.01, 1.14)		-	-
CLA-NIT-PPI-TET	1.65 (0.24, 14.91)	2.16 (0.09, 58.22)	1.80 (0.09, 41.18)	0.04 (0.00, 1.96)	0.38 (0.04, 4.17)	0.23 (0.01, 5.01)	0.38 (0.03, 5.96)	3.79 (0.28, 75.40)		-
NIT-PPI-QUI	2.63 (0.96, 8.09)	3.36 (0.24, 51.22)	2.80 (0.25, 33.01)	0.06 (0.00, 1.91)	0.61 (0.13, 2.66)	0.37 (0.03, 4.32)	0.61 (0.08, 4.58)	5.89 (0.77, 65.73)	1.61 (0.14, 14.33)	

Values given are odds ratios.
The segment below and to the left of the shaded diagonal is derived from the network meta-analysis, reflecting direct and indirect evidence of treatment effects (row versus column). The point estimate reflects the mean of the posterior distribution, and numbers in parentheses are 95% credible intervals.
The segment above and to the right of the shaded diagonal gives pooled direct evidence (random-effects pairwise meta-analysis), where available (column versus row). Numbers in parentheses are 95% confidence intervals.



Values greater than 1 favour BIS-NIT-PPI-TET; values less than 1 favour the comparator treatment. Solid error bars are 95% credible intervals while dashed error bars are 95% confidence interval.

Figure 22: Network meta-analysis of second-line *H pylori* treatment – adherence to treatment – relative effect of all options compared with placebo

Table 33: Network meta-analysis of second-line *H pylori* treatment – adherence to treatment – rankings for each comparator

	Probability best	Median rank (95%CrI)
BIS-PPI-QUI-TET	0.523	1 (1, 6)
BIS-NIT-PPI-TET	0.040	3 (1, 5)
AMO-BIS-NIT-PPI	0.249	3 (1, 9)
AMO-BIS-PPI-TET	0.093	3 (1, 9)
CLA-NIT-PPI-TET	0.077	5 (1, 9)
NIT-PPI-QUI	0.003	6 (3, 9)
BIS-H_2RA-NIT-TET	0.006	7 (3, 10)
AMO-PPI-QUI	0.000	7 (5, 9)
AMO-PPI-TET	0.007	9 (3, 10)
AMO-NIT-PPI	0.002	10 (5, 10)

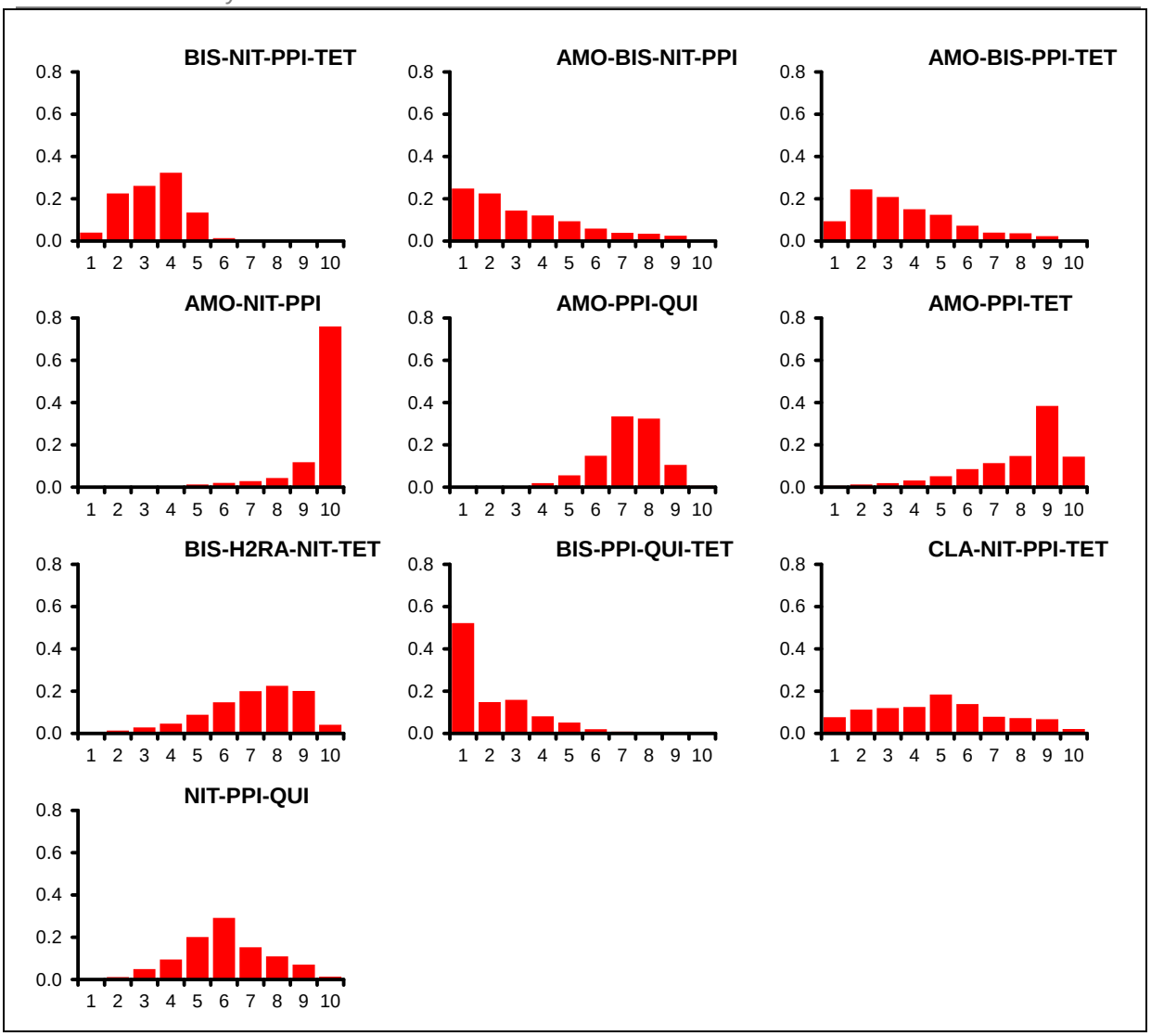


Figure 23: Network meta-analysis of second-line *H pylori* treatment – adherence to treatment – rank probability histograms

Table 34: Network meta-analysis of second-line *H pylori* treatment – adherence to treatment – model fit statistics

Residual deviance	Dbar	Dhat	pD	DIC	Tau
21.63 (compared to 22 datapoints)	86.595	67.218	19.378	105.973	n/a (fixed-effects model)

E.5 *H pylori* second-line eradication by resistance status

This appendix presents the unanalysed data for the outcome eradication by antibiotic resistance status which was considered an important outcome for the following review question:

- What *H pylori* eradication regimens should be offered as second-line treatments when first-line treatments fail?

It was not possible to pool and analyse the data for this outcome due to the following issues:

- Several of the studies measured different antibiotic resistance phenotypes in each trial arm
- Some studies measured resistance to an antibiotic, for example clarithromycin, even though the regimen did not include this antibiotic
- As most studies measured resistance to more than one antibiotic in each arm it is not clear if individuals can be in more than one category and therefore counted more than once

Due to the reasons outlined above the raw data was presented to the GDG in a summary table (below) and was considered as supporting evidence for the eradication outcome but no evidence statement was written.

Table 35: *H pylori* second-line eradication by resistance status

Study	Regimens	CR	CS	MR	MS	LR	LS	TR	TS	AR	AS	CS/LS	CR/LS	CR/LR	CS/MS	CS/MR	CR/MR
Chi (2003)	PPI/BIS/AMO/NIT	6/11 (55%)	16/26 (62%)	5/15[a] (33%)	17/22[a] (77%)												
	PPI/BIS/AMO/TET	8/11 (73%)	23/26 (89%)	13/16[a] (81%)	18/21 (86%)												
Chuah (2012)	PPI/AMO/QUI					2/4[b] (50%)	9/13[b] (69%)				11/17 (65%)						
	PPI/AMO/TET								9/15 (60%)		9/15 (60%)						
Matsumoto (2005)	PPI/AMO/QUI											3/4 (75%)	6/10 (60%)	1/2 (50%)			
	PPI/AMO/NIT														8/8 (100%)	1/1 (100%)	8/8 (100%)
Ueki (2009)	PPI/AMO/CLA/NIT	37/40 (92.5%)															
	PPI/AMO/NIT	35/42 (83%)															
Wu (2011)	PPI/BIS/AMO/TET							0/1 (0%)	16/24 (67%)		16/25 (64%)						
	PPI/BIS/NIT/TET			13/15 (87%)	11/15 (73%)				24/30 (80%)								
Wu (2006)	PPI/BIS/NIT/TET			8/12[c] (67%)	9/9[c] (100%)												
	PPI/CLA/NIT/TET	12/16[d] (75%)	4/7[d] (57%)	7/10[d] (70%)	9/13[d] (69%)												

Clarithromycin resistant (CR); Clarithromycin susceptible (CS); Metronidazole resistant (MR); Metronidazole susceptible (MS); Levofloxacin resistant (LR); Levofloxacin susceptible (LS); Tetracycline resistant (TR); Tetracycline susceptible (TS); Amoxicillin resistant (AR); Amoxicillin susceptible (AS)

N.B. all regimens including NIT used metronidazole as the nitroimidazole; all regimens including QUI used levofloxacin as the quinolone.

[a] *33.3% vs. 73.3% $p < 0.05$; 33.3% vs. 81.3% $p < 0.05$*

[b] *50% vs. 69% N/S*

[c] *67% vs. 100% $p = 0.05$*

[d] *57% vs. 75%; 70% vs. 69% N/S*

E.6 References

Dias, S., Welton, N.J., Sutton, A.J. & Ades, A.E. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. 2011; last updated 16 March 2013; available from http://www.nicedsu.org.uk.

Dias, S., Welton, N.J., Sutton, A.J., Caldwell, D.M., Lu, G. & Ades, A.E. NICE DSU Technical Support Document 4: Inconsistency in Networks of Evidence Based on Randomised Controlled Trials. 2011; last updated April 2012; available from http://www.nicedsu.org.uk.

Mueller, P.S., Montori, V.M., Bassler, D., Koenig, B.A., Guyatt, G.H. Ethical issues in stopping randomised trials early because of apparent benefit. 2007; Annals of Internal Medicine 146 (12): 878-881.