Appendix L Validation of efficacy dataset used in health economic model

Validity of relying on subset of studies reporting dichotomous measures of pain relief

Only some of the included studies report pain relief data in the format required by the health economic model (that is, proportion achieving 30% pain relief and/or proportion achieving 50% pain relief). Because of this, it is possible that an unrepresentative estimate of effect is relied on. This may be a particular concern for treatments with an older evidence-base: the reporting of 30% and 50% pain relief has become more common in recent years.

To investigate the possible impact of this issue, a series of analyses were performed taking advantage of a known relationship between dichotomous and continuous data. Odds ratios may be approximated from continuous data using the formula

$$\ln(OR) = SMD\frac{\pi}{\sqrt{3}},$$

where SMD indicates the standardised mean difference between arms, calculated as

$$SMD = \frac{m_1 - m_2}{s},$$

where m_1 and m_2 are the mean changes in pain from baseline to follow-up in arms 1 and 2 of each study, and s denotes the pooled standard deviation across both groups:

$$s = \sqrt{\frac{(n_1 - 1)SD_1^2 + (n_2 - 1)SD_2^2}{(n_1 + n_2) - 2}},$$

where n represents the numbers of particpants in each arm and SD the standard deviations of the mean changes (see Cochrane Handbook sections 9.2.3.2 and 9.4.6 for details).

This method relies on assumptions that are unlikely to be strictly applicable in the dataset at hand – most notably, that the underlying data follow a logistic distribution, where pain data are known to be more idiosyncratically distributed (see, for example, Moore et al., 2005). For this reason, it was not deemed appropriate to use continuous data to calculate odds ratios that could be directly used in the health economic model. However, a less formal analysis comparing the 2 types of data is useful to investigate the representativeness of data used in the model.

Accordingly, odds ratios (ORs) were approximated from the (larger) pool of studies reporting continuous measures of pain relief and compared with the directly reported odds ratios for 30% or 50% pain relief for the same comparisons. For each pair of treatments for which at least 1 study was available in each category, a stratified fixed-effects meta-analysis was undertaken to compare the pooled ORs estimated in each way and, most importantly, to explore evidence of heterogeneity between strata. If the set of trials reporting dichotomous pain relief does not agree with the results seen in the continuous studies, significant heterogeneity will be detected, suggesting the dichotomous data (and, by extension, the data used in the health economic model) may be a biased sample of the available evidence. See Table L1.

Table L1 Empirical and approximated odds ratios for all comparisons for which dichotomous and continuous pain data are available – exploratory fixed-effects[a] stratified meta-analyses with quantification of heterogeneity

	Dichotomous (30% or 50% pain relief[b,c])				Continuous (mean change in pain[c])				Inter-stratum
Comparison	**N**	**OR**	**(95%CI)**	**Heterogeneity**	**N**	**OR(SMD)[d]**	**(95%CI)**	**Heterogeneity**	**Heterogeneity**
placebo -v- amitriptyline	1	2.04	(0.70, 5.95)	N/A	10	2.16	(1.46, 3.19)	Q=14.76; p=0.098; I^2=39.0%	Q=0.01; p=0.927; I^2=0.0%
placebo -v- cannabis extract	1	2.00	(0.81, 4.96)	N/A	3	1.88	(1.16, 3.06)	Q=2.53; p=0.282; I^2=21.0%	Q=0.01; p=0.906; I^2=0.0%
placebo -v- capsaicin cream	2	6.56	(1.69, 25.44)	Q=1.38; p=0.240; I^2=27.6%	2	2.57	(0.84, 7.92)	Q=0.43; p=0.514; I^2=0.0%	Q=1.08; p=0.298; I^2=7.8%
placebo -v- duloxetine	5	2.43	(1.96, 3.00)	Q=4.27; p=0.371; I^2=6.3%	5	2.34	(1.92, 2.87)	Q=1.50; p=0.826; I^2=0.0%	Q=0.05; p=0.816; I^2=0.0%
placebo -v- gabapentin	3	2.33	(1.50, 3.61)	Q=2.38; p=0.305; I^2=15.8%	13	2.73	(2.21, 3.37)	Q=79.05; p<0.001; I^2=84.8%	Q=0.42; p=0.517; I^2=0.0%
placebo -v- lacosamide	4	1.60	(1.20, 2.13)	Q=0.19; p=0.979; I^2=0.0%	1	1.88	(0.97, 3.62)	N/A	Q=0.20; p=0.659; I^2=0.0%
placebo -v- lamotrigine	6	1.40	(1.04, 1.87)	Q=10.08; p=0.073; I^2=50.4%	3	2.22	(1.34, 3.68)	Q=19.80; p<0.001; I^2=89.9%	Q=2.40; p=0.121; I^2=58.3%
placebo -v- levetiracetam	1	0.73	(0.15, 3.51)	N/A	3	1.37	(0.75, 2.52)	Q=5.97; p=0.051; I^2=66.5%	Q=0.55; p=0.459; I^2=0.0%
placebo -v- morphine	2	3.09	(1.49, 6.41)	Q=0.67; p=0.412; I^2=0.0%	3	2.02	(1.17, 3.49)	Q=2.61; p=0.271; I^2=23.4%	Q=0.83; p=0.361; I^2=0.0%
placebo -v- oxcarbazepine	1	2.04	(1.03, 4.05)	N/A	1	1.92	(1.06, 3.47)	N/A	Q=0.02; p=0.893; I^2=0.0%
placebo -v- pregabalin	14	2.24	(1.92, 2.62)	Q=39.37; p<0.001; I^2=67.0%	10	2.26	(1.92, 2.66)	Q=26.05; p=0.002; I^2=65.4%	Q=0.00; p=0.950; I^2=0.0%
placebo -v- topiramate	1	1.81	(1.12, 2.91)	N/A	4	1.22	(1.01, 1.49)	Q=4.71; p=0.195; I^2=36.2%	Q=2.17; p=0.141; I^2=53.9%
placebo -v- tramadol	2	2.55	(1.49, 4.39)	Q=1.70; p=0.192; I^2=41.3%	3	2.61	(1.74, 3.90)	Q=2.45; p=0.293; I^2=18.5%	Q=0.00; p=0.952; I^2=0.0%
placebo -v- venlafaxine	1	1.77	(1.02, 3.08)	N/A	1	2.13	(0.87, 5.23)	N/A	Q=0.12; p=0.732; I^2=0.0%
amitriptyline -v- gabapentin	1	0.49	(0.17, 1.42)	N/A	3	0.61	(0.30, 1.22)	Q=2.22; p=0.330; I^2=9.9%	Q=0.11; p=0.745; I^2=0.0%
amitriptyline -v- pregabalin	1	1.68	(0.74, 3.82)	N/A	1	4.02	(1.55, 10.42)	N/A	Q=1.85; p=0.174; I^2=45.9%
gabapentin -v- nortriptyline	1	1.29	(0.42, 3.95)	N/A	2	1.26	(0.73, 2.20)	Q=0.00; p=0.948; I^2=0.0%	Q=0.00; p=0.978; I^2=0.0%

[a] fixed-effects meta-analyses were used because, although random-effects models are used elsewhere in this analysis, it is inappropriate to estimate inter-stratum heterogeneity in a random-effects model

[b] to avoid double-counting issues, it was necessary to rely on one or other of 30% and 50% pain relief measures (which, in any case, approximate each other very closely); the 30% measure was preferred, where available, but the 50% measure was used instead if it was the only one reported

[c] where multiple arms in the same RCT addressed the same treatment, data from each were pooled to form a meta-arm for comparison with its common comparator; for dichotomous measures, the numbers and events in the separate arms can simply be summed; for continuous measures, a weighted mean of the relevant means and a pooled estimate of their variances were calculated and used to calculate the SMD

[d] odds ratios approximated from continuous data using standardised mean differences (see above)

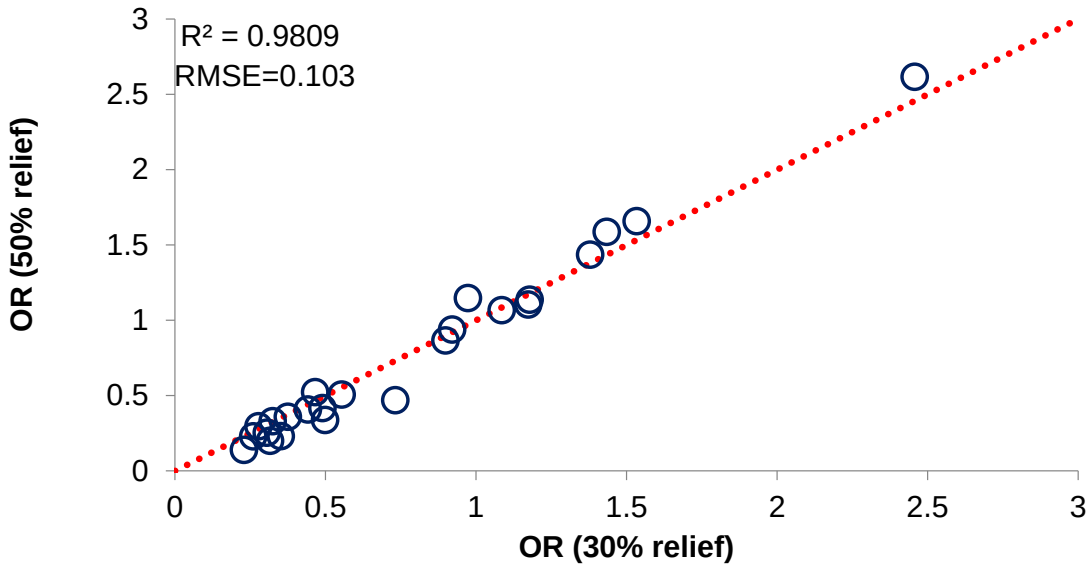
A good degree of homogeneity was seen throughout the comparisons: none of the stratified analyses showed differences between dichotomous and continuous data that that would be considered significant by conventional standards (in tests of heterogeneity, p-values less than 0.1 are often seen as suggestive of non-random differences between strata). The greatest evidence for heterogeneity is for lamotrigine and topiramate: in comparisons with placebo, the former appears somewhat more effective according to the continuous data (suggesting the model inputs may potentially underestimate its efficacy), while the latter has the opposite relationship (so the model may overestimate its efficacy). The head-to-head comparison between amitriptyline and pregabalin also appears to imply some difference between evidence types. However, in none of these cases are the differences of a magnitude that cannot be explained by sampling error.

It is reassuring that treatments with older evidence do not appear to be systematically disadvantaged. In particular, while there are only 3 placebo-controlled RCTs of gabapentin contributing dichotomous evidence to the dataset, it is closely comparable to the OR approximated from the pooled continuous data (which means that the mean effect will be adequately estimated in the health economic model, although uncertainty will be much greater than would be the case if it were possible to derive an entirely robust estimate of effect from all trials). Finally, it is notable that the estimate of capsaicin cream's efficacy derived from dichotomous data is markedly higher than that approximated from the pooled continuous data. However, confidence intervals are broad and, once more, a null hypothesis of homogeneous effects cannot be rejected.

Validity of treating relative measures of 30% and 50% pain relief as a single parameter

As explained in appendix D, the model used to synthesise categorical evidence of pain relief uses single parameters to estimate the relative efficacy of treatments – that is, the extent to which a treatment is better at achieving 30% pain relief than its comparator is assumed to be identical to the extent to which it is better at achieving 50% pain relief. This assumption can be tested by comparing evidence from those studies that have reported both 30% and 50% pain relief, to assess the level of agreement between these data. Figure L1

shows the relationship between estimates of 30% relief and 50% relief from each study that reports both for one or more pairwise comparison.



Dotted line shows equivalence (x=y; OR[30% relief] = OR[50% relief]).

Figure L1 Comparison of odds ratios for 30% pain relief and 50% pain relief

It can be seen that there is almost total agreement between measures of relative effect at the 30% and 50% level. This can be seen as strong validation of the approach taken in the categorical synthesis model.