Table S1. The detailed GRADE profile

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Epidural analgesia	Opiod/no analgesia	Relative (95% CI)	Absolute		
Duration of second stage of labour (Better indicated by lower values)												
8	randomised trials	no serious risk of bias	very serious[1]	no serious indirectness	no serious imprecision	none	732	713	-	MD 5.71 higher (6.41 lower to 17.83 higher)	LOW	CRITICAL
Instrumental birth rate												
8	randomised trials	no serious risk of bias	serious [2]	no serious indirectness	no serious imprecision	none	109/721(-15.1%)	67/721(-9.3%)	RR 1.52 (0.97 to 2.4)	48 more per 1000 (from 3 fewer to 130 more)	MODERATE	CRITICAL
								10.50%		55 more per 1000 (from 3 fewer to 147 more)		
Duration of first stage of labour (Better indicated by lower values)												
4	randomised trials	no serious risk of bias	No serious inconsistency	no serious indirectness	serious[3]	none	230	208	-	MD 17.34 higher (5.89 lower to 40.56 higher)	MODERATE	IMPORTANT
Caesarean section rate												
9	randomised trials	no serious risk of bias	serous[4]	no serious indirectness	serious[3]	none	80/840（9.5%）	100/841（-11.9%）	RR 0.8(0.6 to 1.05)	24 fewer per 1000 (from 48 fewer to 6 more)	LOW	IMPORTANT
								9.20%		18 fewer per 1000 (from 37 fewer to 5 more)		
Spontaneous vaginal delivery rate												
6	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	533/724 (73.6%)	551/732 (75.3%)	RR 0.98 (0.91 to 1.06)	15 fewer per 1000 (from 68 fewer to 45 more)	HIGH	IMPORTANT
								70.30%		14 fewer per 1000 (from 63 fewer to 42 more)		

[1] One outcome showed to favour epidural, three tials showed to favour non-epidural, the others showed there was no significant difference on the second stage of duration.
[2] One outcome showed to favour non-epidural, the others showed there was no significant difference on it. [3] The CI was rather wide. [4] One outcome showed to favour epidural, the others showed there was no significant difference on it.

Criteria for assigning grade of evidence:

Type of evidence Randomised trial = high **;** observational study = low**;** any other evidence = very low

Decrease grade if: Serious (-1) or very serious (-2) limitation to study quality; important inconsistency (-1)**;** some (-1) or major (-2) uncertainty about directness**;** imprecise or sparse data (-1)**;** high probability of reporting bias (-1)

Increase grade if: Strong evidence of association-significant relative risk of > 2 (<0.5) based on consistent evidence from two or more observational studies, with no plausible confounders (+1)**;** very strong evidence of association-significant relative risk of > 5 (< 0.2) based on direct evidence with no major threats to validity (+2)**;** evidence of a dose response gradient (+1) **;**all plausible confounders would have reduced the effect (+1)

Definitions of grades of evidence:

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.