Appendix 3 (as submitted by the authors): Quality assessment criteria

	High quality	Medium quality	Low quality
1. Selection bias:	Studies with randomisation, allocation concealment, similarity of groups at baseline	RCTs with some deficiencies in randomisation e.g. lack of allocation concealment, or non-randomised studies with either similarities at baseline or use of statistical methods to adjust for any baseline differences	Non randomised, with obvious differences at baseline, and without analytical adjustment for these differences.
2. Performance bias:*	Differed only in intervention, which was adhered to without contamination, groups were similar for co-interventions or statistical adjustment was made for any differences	Confounding was possible but some adjustment was made in the analysis	Intervention was not easily ascertained or groups were treated unequally other than for intervention or there was non-adherence, contamination or dissimilarities in groups and no adjustments made.
3. Measurement bias:	Outcome measured equally in both groups, with adequate length of follow-up (i.e. at least 6 weeks postpartum), direct verification of outcome, with data to allow calculation of precision estimates.	Inadequate length of follow up or length not given	Inadequate reporting or verification of maternal mortality or differences in measurement in both groups
4. Attrition bias:	No systematic differences in withdrawals between groups and with appropriate imputation for missing values		Incomplete follow-up data, not intention-to-treat analysis or lacking reporting on attrition

*Blinding was not a quality assessment as blinding of participants or caregivers to intervention types was not possible