Online supplementary file 4 – Guide for judgement of modified GRADE synthesis

GRADE Domain	Explanation
Study limitations:	the evidence for each stated PF and its related outcome were assigned categories of either: (1) Few limitations, where most QUIPS domains were classed as low risk, (2) many limitations, where most QUIPS domains were unclear or moderate risk, or (3) very many limitations, where most QUIPS domains were high risk.
Consistency of results:	related to variability in results when two or more studies investigated the same PF and outcome, categorised as yes or no. Where only one study investigated a PF this was not applicable.
Effect sizes:	
	were judged as small, moderate or large for each factor (see Table 2 footnote for threshold values). Moderate to large effect sizes increase the likelihood of a true association ²⁶ dependent on sample size.
Precision of results:	evaluated by interpretation of confidence interval width, identification of sample size calculations to establish study power and the number of studies that investigated each PF, as inadequate study numbers or participants introduce imprecision. Based on these evidence was categorised as precise or imprecise.
Publication bias:	was considered present in all factors unless investigated by multiple cohort studies. ²⁶
Overall quality:	Using the above an overall quality grade, ranging from very low quality to high quality, was assigned to each PF using the process described by Huguet et al. ²⁶