Appendix 2. Quality assessment criteria

Studies were considered "unlikely to inform the primary question" if they had one or more of the following flaws:

- No adjustment for any measure of condom use unless authors report trivial differences comparing estimates from models including and not including condom use, or
- Unclear measurement of exposure, including one or more of the following:
    - Failure to include time-varying analysis of exposure, if appropriate (e.g., time-varying analysis may not be necessary for studies with extremely short follow-up periods).
    - It was unclear what methods were used in the comparison group, especially in relation to hormonal contraceptive use; or, if such usage was detailed, specific risk estimates were not presented for comparisons between hormonal contraceptives and Cu-IUD use*.
    - The interval of time between study visits ("intersurvey interval") was longer than 6 months, with contraceptive use measured only at each interval endpoint (and thus providing only limited information about possible contraceptive switching during the intersurvey interval). (Note: if variation in length of intersurvey interval occurred within an individual study, such that some intervals were 6 months or less and other intervals were longer than 6 months, we included only data from intervals that were 6 months or less).

Studies considered "informative but with important limitations" had none of the flaws described above.

*In this review we have assumed that where a study does not specify what type of IUD was used, that most, if not all, usage related to Cu-IUDs rather than LNG-IUSs, reflecting patterns of IUD use in the countries studied.

In addition, the following criteria apply to RCTs:

1. The previous criteria apply equally to RCTs as well as observational studies
2. Given that RCTs have the potential to have fewer biases and opportunity for confounding, we considered whether such studies might be classified as "informative with few limitations". We decided that such studies might be classified in this way, if most (if not all) of the following elements were present:
    a. Publication of a full trial protocol, including analysis plan, prior to conduct of the study.
    b. Good randomization procedures and completion (clear description of randomization procedures, with good concealment of allocations at point of assignment (ideally via a remote randomization procedure) and, where appropriate, a process which included stratification on key prognostic variables.
    c. Good adherence to allocated treatments, with limited discontinuation or change of method of contraception. If occurs in less than 20% of trial participants (and most other criteria are met) may still be deemed "informative but with few limitations"; if more than 20% affected but

investigators adjusted appropriately for this in analysis may be "informative with important limitations"; if more than 20% affected and no appropriate adjustment made, trial will be deemed "unlikely to inform the primary question".

d. Blinding of participants and study personnel to allocated treatments.

e. Independently ascertained outcomes, with personnel responsible for ascertaining the main trial outcomes blind to allocated treatments.

f. High rates ($\geq$ 80%) of follow up of trial participants resulting in high rates of ascertainment of outcomes.

g. Analyses conducted blind to treatment allocation, with time varying analysis of key confounders, including estimates of condom use during the study; with intention to treat analyses the primary comparison between groups.

h. Results reported as per pre-defined analysis plan, with subgroup analyses predefined or clearly designated as post hoc analyses (and with a clear justification for their conduct).

It was recognised that many trials in reproductive health care will not fulfil all of these criteria (for example, trials of different contraceptive methods are likely to be open, with participants and clinicians responsible for their routine care unblinded to the treatments allocated). Nonetheless, clarity about how the researchers sought to minimise bias and confounding after treatment allocation should enable a judgement to be made about the extent to which evidence from the trial was limited. It is possible, although probably highly unlikely, that an observational study purposively established to investigate the possible association between hormonal contraception and HIV acquisition in users meets most of the above criteria. In such a situation a judgement will be made on whether that study was 'informative with few limitations'.