

Table E. Operationalization of scoring the Menstrual Disorders and Subfertility Group external validity quality items* [posted as supplied by author]

External validity items	Scoring of item	Notes about scoring
1. Were the inclusion and exclusion criteria for entry clearly defined?	The reviewer scores the item separately on whether the inclusion criteria and the exclusion criteria were explicitly reported.	Reviewer provides 2 separate assessments, the first for the inclusion criteria clearly defined (Y/N) and the second for the exclusion criteria clearly defined (Y/N).
2. Were the outcome measures used clearly defined and reported?	The reviewer considers the clinical pregnancy outcome to be clearly defined and reported if the RCT authors reported how clinical pregnancy was defined (e.g. fetal heart activity on ultrasound) AND when it was assessed (e.g. 6 weeks after embryo transfer). Similarly, ongoing pregnancy was considered to be clearly defined and reported if the authors reported how ongoing pregnancy was defined (e.g. viable intrauterine fetus) AND when it was assessed (e.g. beyond 12 weeks gestation).	To earn a 'Y' for this item, RCT authors needed to report how <u>both</u> of the outcomes they assess were defined <u>and</u> at what time.
3. Were the accuracy, precision, and observer variation of the outcome measures adequate?	The reviewer considers the accuracy and precision of the outcome measures to be adequate if there is a clinical confirmation of pregnancy by ultrasound.	
4. Was the timing of the outcome measures appropriate?	The timing of the outcome must be reported and it must be appropriate.	To score a 'Y', the RCT authors must report the timing of each of the outcomes that they assess as also with item 2 above, but in addition, the timing of each outcome must be considered appropriate by the reviewer,
5. Was a power calculation performed? †	The reviewer determines if enough information is reported in order to score a 'Y'.	

*The Cochrane Menstrual Disorders and Subfertility Group module[Clarke et al 2007, The Cochrane Library] does not provide an operationalization of the scoring for these external validity items. Therefore, the operationalization of the scoring of the items described above was developed by the author (EM). Each item is scored as 'Y' for 'Clearly Yes', '?' for 'Not sure' (i.e. cannot be determined based on the text), and 'N' for 'Clearly No'. RCT investigators were asked for clarification for the items scored as 'Not sure'.

†Item E5 is not included on the Cochrane Menstrual Disorders and Subfertility Group scale but has been included in other published quality evaluation scales, and therefore this item was included and extracted for this review to ensure extraction of all elements related to external validity.