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

Internal validity items	Scoring for item	Notes about scoring
1. Was the method of randomization adequate? †	A random (unpredictable) assignment sequence. Examples of adequate methods are computer generated random number table and use of sealed opaque envelopes. Methods of allocation using date of birth, date of admission, hospital numbers, or alternation should not be regarded as appropriate.	If the text states that patients were randomly assigned but provides no further details, the item was scored as '?'.
2. Was the assigned treatment adequately concealed prior to allocation?	Clearly yes Score 'Y' ·Some form of centralised randomisation scheme, such as having to provide participant details by phone to receive treatment group allocation ·A scheme controlled by a pharmacy ·In a pharmaceutical study, sequential administration of pre-numbered or coded containers to enrolled participants ·An on-site computer system, given that allocations are in a locked unreadable file which can be accessed only after inputting participant details ·Assignment envelopes, provided that they are sequentially numbered, sealed, and opaque ·Other combinations which appear to provide assurance of adequate concealment ·Unclear Score '?' ·Assignment envelopes, without description of adequate safeguards ·Use of a "list" or "table" ·Flip of a coin ·A trial in which the description suggests adequate concealment, but other features are suspicious - for example, markedly unequal controls and trial groups ·Stated random, but unable to obtain further details ·Clearly no Score 'N' ·Alternation	

Internal validity items	Scoring for item	Notes about scoring
	·Case record numbers, dates of birth, day of the week, or any other such approach ·Any allocation procedure transparent before assignment, such as an open list of random numbers	
3. Were the outcomes of patients who withdrew or were excluded after allocation described and included in an 'intention to treat' (ITT) analysis?	For ITT analysis, all randomized patients are reported/analyzed in the group they were allocated to by randomization for the most important moments of effect measurement (minus missing values) irrespective of noncompliance and cointerventions.	If ITT was not used in the trial analysis, but the necessary data were reported for an ITT analysis for the meta-analysis , the item was scored as A.
4. Were the outcome assessors blind to assignment status?	The reviewer determines if enough information about the blinding is given in order to score a 'Y'.	
5. Were the treatment and control group comparable at entry?	The reviewer determines if enough information about the similarity of the groups at baseline is given in order to receive a 'Y'.	
6. Were the subjects blind to assignment status following allocation?	The reviewer determines if enough information about the blinding is given in order to score a 'Y'.	
7.a. Were the acupuncturists blind to assignment status?	The reviewer determines if enough information about the blinding is given in order to score a 'Y'.	
7.b. Were the IVF procedure physicians blind to assignment status?	The reviewer determines if enough information about the blinding is given in order to score a 'Y'.	
8. Were the care programs, other than the trial options, identical (i.e. was there a cointervention)?	Cointerventions should either be avoided in the trial design or similar between the index and control groups.	
9. Were the reasons for withdrawals stated? †	The number of participants who were included in the study but did not complete the observation period or were not included in the analysis must be described and reasons given.	
10. Was the percentage of dropouts less than 10%?	If the percentage of withdrawals and drop-outs does not exceed 10% for the primary outcome and does not lead to substantial bias a 'Y' is scored.	

Internal validity items	Scoring for item	Notes about scoring
11. Was there sham credibility testing? †	For trials that used a sham control, this item is scored as 'Y' if there was an attempt to confirm patient blinding by asking patients about their treatment assignment.	

*The Cochrane Menstrual Disorders and Subfertility Group module[Clarke et al 2007, The Cochrane Library] provides an explicit operationalization of the scoring only for the allocation concealment item (i.e. item 2). Therefore, for operationalizing the scoring of the other items, we used the operationalization of scoring developed and used by the Cochrane Back Review Group.[van Tulder et al 2003 Spine 28:1290-9] 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'. For the items scored as 'Not sure' based on the text alone, the authors were contacted and asked for clarification and/or further explication. If the additional information provided by the author clarified the methods, then the item initially scored as '?' was changed to either 'Y' or 'N'. For calculating summed scores, items scored as 'Y' were counted as 1and items scored as '?' (i.e. still '?' after contacting authors) or 'N' were counted as 0.

†Items I1, I9, and I11 are not included on the Cochrane Menstrual Disorders checklist, but have been included in other published quality evaluation scales, and therefore were included and extracted for this review to ensure extraction of all elements related to internal validity.