

中华医学会生殖医学分会（CSRM）指南共识制定规范

中华医学会生殖医学分会

一、定义

指南共识类文章是指具有学术权威性的指导类文章，包括指南、规范、专家共识、草案等^[1]。指南共识的制定，是通过系统评价对相关证据进行再评估和整合，并且考虑到利/弊后，经讨论达到一致性意见或提出最恰当的推荐意见并形成具有权威性的医疗文件，以帮助患者和医护人员做出临床诊疗重要决策。而实际工作中，指南共识并不适用于所有情况和患者，最终的诊疗决策是根据患者意愿、具体病情、医生认知等综合影响因素而制定。

常见的相关名词：

临床治疗管理办法

由国家卫计委根据《执业医师法》、《医疗机构管理条例》、《医疗事故处理条例》等有关法律、法规和规章制定并颁布的一些管理办法^[2]，其目的是加强医疗临床管理，建立医疗管理制度，促进医学科学发展和医疗技术进步，提高医疗质量，保障医疗安全。

临床诊疗规范

为加强医疗机构在疾病诊断和治疗方面的规范化管理，提高医疗质量，保证医疗安全，专家们根据《执业医师法》、《医疗机构管理条例》、《医疗技术临床应用办法》等有关法律、法规和规章制度，深入讨论制定的规范^[2]。有助于全国各地、各级医疗机构广大医生诊疗行为的标准化、科学化和规范化，有效提高临床医生的诊疗水平。

专家共识

专家共识是专家们对严谨客观的临床试验进行客观分析，同时对临床医生或相关专家的经验、建议和意见等主观性经验进行规范化研究分析，经讨论达成专家一致性共识或意见，是对疾病预防与治疗的方向性指导^[2]。

临床实践指南（clinical practice guideline, CPG）

临床实践指南是在特定的临床环境下，针对特定的临床问题，经系统研究后制定发布，帮助医务人员和患者对患者本身的健康问题做出恰当的认识和处理提供简

要明了、通俗易懂的指导性意见（推荐建议）^[3]。指南是经过长时间循证医学依据论证和有关理论知识及临床经验总结出来的指导性意见，具有权威性、宏观性、适用范围广等特点。

二、目的

CSRMM 指南共识主要目的是为中国人类生殖和胚胎学领域提供临床推荐意见，以推动其发展。CSRMM 指南共识是在已有的证据基础之上，对一个特定的临床问题（如男性不育，PCOS，PGD/PGS），根据已有的临床研究/证据形成专家共识或对证据评价形成具有权威性的临床实践指南。CSRMM 指南共识可以用于临床医务人员的培训，指导临床工作，促进患者和医务人员的交流，以及为中国一些医疗政策的制定提供参考。

该手册是为了 CSRMM 更加规范、客观、详细、准确制定指南共识而编写，以保证指南共识制定的科学性、公正性和权威性，对中国生殖医学的良性发展起到积极影响。该手册的编写基于国际指南编写方法学、国际方法质量评定标准的相关文献，借鉴了国外指南共识手册的编写经验。

三、临床诊疗规范制定过程

临床诊疗规范是受国家职能部门委托，经过 CSRMM 常委会讨论，组成编写小组，根据《执业医师法》、《医疗机构管理条例》、《医疗技术临床应用办法》和《人类辅助生殖技术和人类精子库管理办法》等有关法律、法规和规章制度，系统总结相关疾病或技术的标准化诊断、治疗、操作方法，以及近年来发展的新成果和新技术。编写小组形成初稿后，征求委员的意见，收集意见后在常委会上进行讨论定稿。

四、CSRMM 专家共识制定方法和过程

制定高质量专家共识是国内外实施规范化医疗、提高医疗质量和降低医疗投入十分有效的方法。

（一）专家共识制定方法分为非正式和正式的专家共识制定方法：

1. 非正式的专家共识制定方法，较为简单，专家共识制定者组织专家通过会议形式，将会议讨论后达成的共识意见形成推荐指南即专家共识，由专业学会或政府机构进行发布。这种指南的制定缺乏相关的证据基础，推荐意见易受专家个人经验和主观判断的影响，专家推荐的干预方式在实际的临床应用中并不能确保真正对患者有利，因此，这种指南的质量一般较差，不具备真正意义的可靠性。如，系统综述问题的确定、证据的解释、建议的制定等。

2. 正式的专家共识制定方法采用特尔菲法（delphi technique）、名义群体法（nominal-group technique）和共识会议法等方法。推荐采用共识会议法，由相关专家、群体、代表等以投票、排序或公开讨论等方式针对相关的研究证据进行评估，最终专家小组成员得出一致性程度较高的推荐意见。正式的专家共识指南制定法虽然回顾了相关的研究证据，但在制定推荐意见时没有相关证据质量评价，专家的主观意见仍占主导地位。

（二）CSRM 专家共识制定的过程

1. 题目的选定和申请

申请人通过填写申请表（见附录 A），向 CSRM 委员会申请启动一个新的专家共识的编写工作。所有递交的申请，都由 CSRM 委员会秘书组收集、整理后，提交 CSRM 常委会，由常委会讨论决定其申请是否可以进入正式的编写流程。

2. 成立编写小组

有关申请经常委会讨论通过后成立编写小组并确定负责人。

1) 人员组成：一旦新的专家共识申请被批准后，负责人需要考虑多元化的编写小组人员，可包括：业内专家、非专家医师、患者、相关的医疗服务人员（卫计委人员）和一名 CSRM 研究专家，每个编写小组成员在 10-15 名最佳，适当考虑地域、性别平衡问题。

2) 职责：在编写的不同阶段，编写负责人将在讨论会议中组织特定培训，目的在于提高并均衡编写小组成员的水平、确定人员分工和职责，如有需要可以组织编写研讨会。如果编写小组成员连续两次不能参会，负责人可以要求其退出小组。如果此时编写正在进行中，通常将不再添加新的组员。如需要另外增加专家或替换组员，则需要通过组内讨论。

为确保编写小组能够有效运行，所有组员须遵守以下职责：接受不同观点，

开放性讨论，证据胜过观点，共同承担责任，定义保密范围，经 CSRM 常委会通过后才可以 CSRM 名义出版，保证参加所有会议，所有组员必须提供关于机密性的公开声明（见附表 B）。

3. 确定范畴和形成关键问题

- 1) 明确目标：明确所有目标，将对哪些患者及目标人群实施，与其他文件的关系。
- 2) 确定范畴：确定适用范围的方法包括初步文献查阅，或者通过咨询目标使用者及患者，并记录其期望及参数清单。咨询的方法包括个人或群体访谈，也可以进行问卷调查。在大致确定应列入的问题，应排除的问题之后，范畴可以根据记录的清单来确定（附 C）。包括专家共识关注的疾病、患者群体、不同类型的干预与治疗措施、主要结局，以及针对的医疗机构。

项目记录包括：

总体目标（例如，成本，有效的医疗服务）

目标使用者（例如，胚胎学家，妇科学家）

患者条件（例如，诊断类型，不孕年限）

患者人群（例如，年龄限制）

医疗保健设置（例如，法律法规、支付制度、二级或三级治疗）

诊断方法（例如，精液分析、腹腔镜检查）

干预/治疗（例如，手术治疗、心理学治疗、配子捐赠、生活保健建议）

治疗结局（例如，妊娠以及并发症率）

患者预后（例如，轻度卵巢过度刺激，双胎妊娠）

与其他文件的关联（例如，其他指南、道德准则和法律）

- 3) 范畴评估：完整有效、中国范围、现实可行。一旦范畴最终确定，编写小组则开始规划关键问题用于文献查阅，同时应设定制定过程的时间流程。

- 4) 关键问题的形成：关键问题应当清晰、重点突出并且紧密的围绕在所讨论的主题范围之内。关键问题不仅在文献系统性回顾中很重要，同时是专家共识制定推荐的方向。常见的关键问题包括 3 类：干预问题、诊断问题和预后问题。在确定关键问题的时候，建议使用 PICO（Patients/population, Interventions, Comparisons, Outcomes）（患者/人群，干预，对照，结局）原则，即一个问题要尽量包含以上

四个要素，一般含 15-20 个关键问题最佳。编写小组通过达成一致意见来决定最终的关键问题。

4. 系统文献检索和识别

证据的甄别和选择是回答关键问题的必不可少的步骤，首先是查阅现存的提出相同问题的指南。其次，关键问题应当作为关键词或检索词的形式，便于通过系统全面检索进行证据查询，避免偏倚。

1) 现有指南共识

查阅是否有现存的指南共识以供参考，推荐应用以下网址查询：

美国生殖医学学会(www.asrm.org)

欧洲人类生殖及胚胎学会(www.eshre.eu)

国际指南网(www.g-i-n.net)

美国国立指南库(www.guideline.gov)

澳大利亚国家卫生和医学研究委员会(www.nhmrc.gov.au)

英国国家卫生署(www.nice.org.uk)

苏格兰院间指南网(www.sign.ac.uk)

世界卫生组织(www.who.int)

2) 文献查阅

文献查阅是制定一个有效专家共识或指南的关键步骤。不完整或有偏倚的文献评估将导致不恰当的推荐意见。

a. 方法：根据关键词，系统全面检索，查找证据，避免偏倚。运用逐步回归法采用最佳证据处理关键问题。规范的搜寻筛选方法大致顺序如下：系统评价/Meta 分析、随机对照研究、非随机对照研究、观察性研究、个案报道/专家意见。**b. 推荐遵循以下规则：**只有发表的同行评议文献才能被考虑在内；除极少数特殊情况外应避免使用摘要；应避免用未发表的临床研究来支持任何推荐。**c. 范围：**文献检索库至少应包括以下几个数据库：MEDLINE/PubMed、Cochrane Library、CNKI 等其他。

5. 形成专家共识初稿

在整个专家共识制定过程中，编写小组成员均应达成共识，共识内容包括提出关键问题、用最佳证据回答关键问题以及形成建议。推荐采用共识会议法，由编写小组组织相关专家、群体、代表等以投票、排序或公开讨论等方式针对相关的研究

证据进行评估，最终专家小组成员得出一致性程度较高的推荐意见，即专家共识的初稿。

五、CSRM 循证实践指南制定的过程

循证实践指南是基于系统全面收集文献资料，严格评价相关证据的质量，使用科学的方法对证据进行再评价，综合考虑证据级别、证据强度和专家组成员的实践经验之后提出更具权威性的推荐意见。推荐意见应与证据质量有明确的联系，循证CPG已成为制定指南的趋势。

CSRM 循证实践指南制定是建立于国际指南编写方法学、国际方法质量评定标准的相关文献的基础之上，纳入了国际指南制定方法学最新证据和可利用的资源，来源于：

- European Society of Human Reproduction and Embryology(ESHRE)
(www.eshre.eu)
- American College of Cardiology Foundation and American Heart Association
(www.americanheart.org)
- European Society of Cardiology (ESC) (www.escardio.org)
- National Health and Medical Research Council (NHMRC) (www.nhmrc.gov.au)
- National Institute for Health and Clinical Excellence (NICE) (www.nice.org.uk)
- New Zealand Guidelines Group (NZGG) (www.nzgg.org.nz)
- Scottish Intercollegiate Guidelines Network (SIGN) (www.sign.ac.uk)
- World Health Organization (WHO) (www.who.int).
- Guidelines International Network (<http://www.g-i-n.net/>)
- Institute of Medicine (<http://www.iom.edu/>)

循证实践指南的具体制定过程如下（1-4 同 CSRM 专家共识制定过程）：

- 1.题目的选定和申请；
- 2.成立编写小组；
- 3.确定范畴和形成关键问题；
- 4.系统文献检索和识别；

5. 证据的总结评价和推荐分级

1) 证据选择: 编写小组制定一个文献/研究的纳入标准, 检索出来的所有文献根据标题、摘要进行初筛, 将不符合标准的研究/文献予以排除。

2) 质量评价: 对所选证据资料进行质量评估是确保指南基于最高等级的可用证据, 形成推荐意见的方法, 每一个研究都应进行其有效性和实用性的质量评估; 按文献研究类型应用不同的调查表(附 D1-D4), 这些调查表以美国卫生保健质量和研究署(AHRQ)所提出的关键领域为基础, 用于科学证据的强度分级。此类调查表也被英国国家临床优化研究所(NICE)和苏格兰学院间指南网络(SIGN)所用。如果一项研究的质量评估为低级别, 那么该研究将被排除; 如果与专家观点相比不能找到更佳的证据, 此研究将作为低级别证据。

质量评价的内容包括: a. 研究发现的可实施性: 也被称为外在效度或可适性, 与关键问题的 PICO 各部分相关。按 PICO 原则, 比较可用文献, 选择拥有相关证据的文章; b. 研究效度: 研究效度应当达到研究设计的目的并可实施, 避免系统性误差或偏倚。医疗试验中可能出现的系统性误差主要包括选择偏倚、实施偏倚、失访偏倚、检出偏倚。

这四种偏倚可评为“发生”、“未发生”、或“不确定”三个级别。用从低危到高危偏倚的范围来总结这四种偏倚。这一效度评估可以做为研究的纳入阈值(如果涉及多个效度准则, 预示着高危偏倚, 应成为排除标准的理由), 也可用于不同研究结果可能的解释。

3) 证据总结: 纳入指南的研究数据需根据标准模板进行提炼(附 E1 和 E2), 按照统一标准进行总结。按 PICO 原则, 总结内容主要包括: 研究纳入人群(样本量, 年龄等); 干预措施(随访年限, 干预类型等); 对照情况(如 IUI 与指导同房相比)和结果(有效性等); 证据表根据国际网络的指南(GIN)证据表格工作小组提出的建议进行更新, 保存完整的证据表格, 以提供给指南使用者。

4) 证据分级: 根据证据的支持力度, 对纳入的证据进行评分。与 SIGN 和 NICE 同样, CSRMM 根据美国牛津循证医学中心的证据分级水平及推荐依据, 使用以下证据分级:

1 级: 来自多个随机试验或 meta 分析的数据

2 级: 来自单个随机试验或大样本非随机研究的数据

3 级：来自回顾性研究、个案研究的数据

4 级：专家意见（包括描述性综述）

以质量评估列表的结果为基础，证据分级用++（高质量），+（中等质量）或-（低质量）来补充，这一证据分级结果如下：

推荐分级	证据分级	研究
A	1++	高质量 meta 分析、RCT 系统回顾或误差风险极低的 RCT
	1+	较好的 meta 分析、系统回顾或低误差风险的 RCT
B	1-	Meta 分析、系统回顾或误差风险较大的 RCT
	2++	高质量的病例对照研究或队列研究的系统回顾 高质量且极低误差风险的病例对照研究或队列研究并且因果关系可能性大
	2+	较好的且低误差风险的病例对照研究或队列研究，因果关系的可能性中等
C	2-	高误差风险的病例对照研究或队列研究且不存在因果关系
D	3	非分析型研究，如病例报道、病例系列
GPP	4	专家意见

A 级:证据极有效，推荐；B 级:证据有效，可推荐，也可能在将来更高质量的新证据出现而改变；C 级:证据在一定条件下有效，应谨慎应用研究结果；D 级:证据的有效性局限，只在较窄的范围内有效。

GRADE 证据水平与推荐级别（2004 年）[5]

等级推荐、制定、发展和评价（The Grading of Recommendations Assessment Development and Evaluation, GRADE）工作小组制定了一个用于系统回顾和指南制定中进行证据质量分级的框架。GRADE 是“结局中枢”：对每一个结局进行分级，在一个研究中同一个结果可能有不同质量分级，并且分级涉及一系列的证据。

在 GRADE 方法中，随机对照试验（RCT）作为高质量证据，观察性研究作为低质量证据来支持干预的预期效果。5 个因素可能导致证据质量级别降低，3 个因素可能使级别升高。对结局证据质量进行分级的 GRADE 方法被归纳为下表：

系列证据来源	原始质量分级	可能降低等级的因素	可能提升等级的因素	最终证据质量*
随机试验	高	1. 误差风险	1. 大的干预效应	高(A)
		2. 矛盾性	2. 剂量反应	中等(B)
观察性研究	低	3. 间接性	3. 潜在的混杂因素将降低已证实的干预效应，	低(C)
		4. 不严密		
		5. 发表偏差	极低(D)	

*证据质量界定:

高: 进一步研究将不大可能改变评价效力的可信度

中等: 进一步的研究将有可能对评价效力可信度产生重要影响并且可能改变评价结果

低: 进一步的研究极有可能对评价效力可信度产生重要影响并且改变评价结果

极低: 任何评价效力都不确定

虽然 GRADE 被广泛应用, 但它仍有一定的局限性: 它只适用于一些特定干预措施的回顾分析和内容推荐; 在指南制定过程中, 一些关键问题包括诊断、风险、预后并不适用这一方法。因此在 CSR 指南中, 在对比已经明确定义结局的干预措施的关键问题时运用 GRADE。在建立证据表格以及对各个研究进行了质量评估之后, 起草 GRADE 纲要 (根据 GRADE 工作小组提出的模板) 并形成指南。

6. 推荐意见的制定

1) 推荐分级: 通过对证据的筛选和评级之后, 要对证据进行总结, 形成简单明了的推荐意见。而在 GRADE 系统推荐级别只分为“强”、“弱”即 1、2 两级。当明确显示干预措施利大于弊或者弊大于利时, 将其列为强推荐(推荐级别 1)。当利弊不确定或无论质量高低的证据均显示利弊相当时, 则视为弱推荐(推荐级别 2)。在指南或者共识中, 推荐意见的标识举例如下: 1A: 基于 A 级证据, 强烈推荐; 1B: 基于中等强度证据, 强烈推荐等。2C: 基于弱证据, 弱推荐; 2D: 基于非常弱的证据, 弱推荐; 在推荐强度分级时, GRADE 除考虑证据质量外, 其它一些因素也会影响推荐意见的强弱, 包括患者价值观和意愿。尽管该证据质量高, 但不同价值观和意愿的患者选择会有不同, 就不能确定为强推荐。另外, 还要考虑资源的合理利用, 推荐措施的获益程度与风险以及成本比。

2) 形成推荐: 需得到全部指南编写小组赞成。措辞和定义必须清楚和标准化, 确保建议之间的一致性和相容性, 并且容易用于临床实践。制定指南时建议的标准措辞如下:

推荐的措辞	
强的建议（1级）	临床医生应该 推荐 表明
	（负面的结果） 临床医生不应该 不推荐/未表明 可能有害
弱的建议（2级）	可能推荐 有条件的推荐 建议 临床医生可能 临床医生可以考虑 临床医生可能考虑
专家建议(GPP)	指南编写小组推荐

7. 指南草稿的撰写

CSRM 指南可以用汉语或汉语和英语书写。指南应该通俗易懂，在不同临床操作环境中能够应用。指南的书写参考格式：

1) 指南的介绍：分为临床意义和指南范围。

- a.临床意义包含患病率、实践中的差异、理想保健的提供、个人和社会成本。
- b.指南范围，包括目的信息，纳入和排除结果，目标人群（如果有需包括定义和分级）。

2) 关键问题（主要是证据和推荐指南之间的关系，是全文的重点）：编写小组针对关键问题制定一个逻辑顺序。其中一个方法是按照病人的路径，从临床评估(例如症状、危险因素)，其次是诊断、治疗方案、后续、并发症、信息的提供。每个关键问题需要从以下几个方面报告：

- a.关键问题本身
- b.解释性文本，总结所选的临床证据
- c.结论部分和注意事项：形成的建议由从证据中得到的主要结论和由指南编写小组讨论的注意事项所组成。源于证据的建议框架应提及可能的注意事项，包括证据的质量、利益和危害、患者价值观、利益相关者的可接受性和资源利用，必要时可以扩展。应记录不同意见的陈述。注意事项的内容还要考虑道德、伦理和法律。

d.分级建议

e.参考文献

3) 总结部分（索引、简写说明、指南编写小组、下一步研究指导、方法学、指南评阅人等）

所有建议的总结（压缩版）将作为指南的一个单独的文档。这个版本包含完整指南里最重要的内容，即所有关键问题和建议。该版本也将在中华医学会生殖医学分会网站上公布。

六. CSRM 指南共识咨询及同行评议

专家共识和循证实践指南草稿需要其未来的读者及制定相关人员的审阅，最后由 CSRM 委员会确定。

所有参与指南共识制定的人员、国内学会组织等都可以对指南共识进行评阅；另外草稿还将公布于 CSRM 网站，会员亦可以进行点评修订。所有参与审阅并发表意见的人都要发表相关利益冲突声明，并在 4 周之内提交意见。也可举行一个开放会议进行审订。

在评议人员的评论处理报告中，评议人员的评论将被制成列表并讨论。指南共识的修改必须由整个编写小组同意并在报告中提及。如果没有变化，需记录原因。CSRM委员会最终批准指南共识。

七. CSRM 指南共识的发布

传播 CSRM 指南共识是编写小组工作的一个延续，它包括指南共识的访问、宣传指南共识的可用性，广泛传播指南共识。只有大力传播和执行指南共识，指南共识才起作用。如果没有起到作用，就是浪费时间、精力和成本。

指南共识的发布渠道多种多样，CSRM 指南共识的常用发布途径：

- 使用简短的总结
- 推广指南共识（CSRM年会、CSRM电子新闻、CSRM网站滚动播出等）
- 发表在专业期刊
- 发布在互联网上和和相关网站建立链接（CSRM网站）

- 使用常见的通信途径
- 会议巡讲

八. CSRM 指南共识的实施和评价

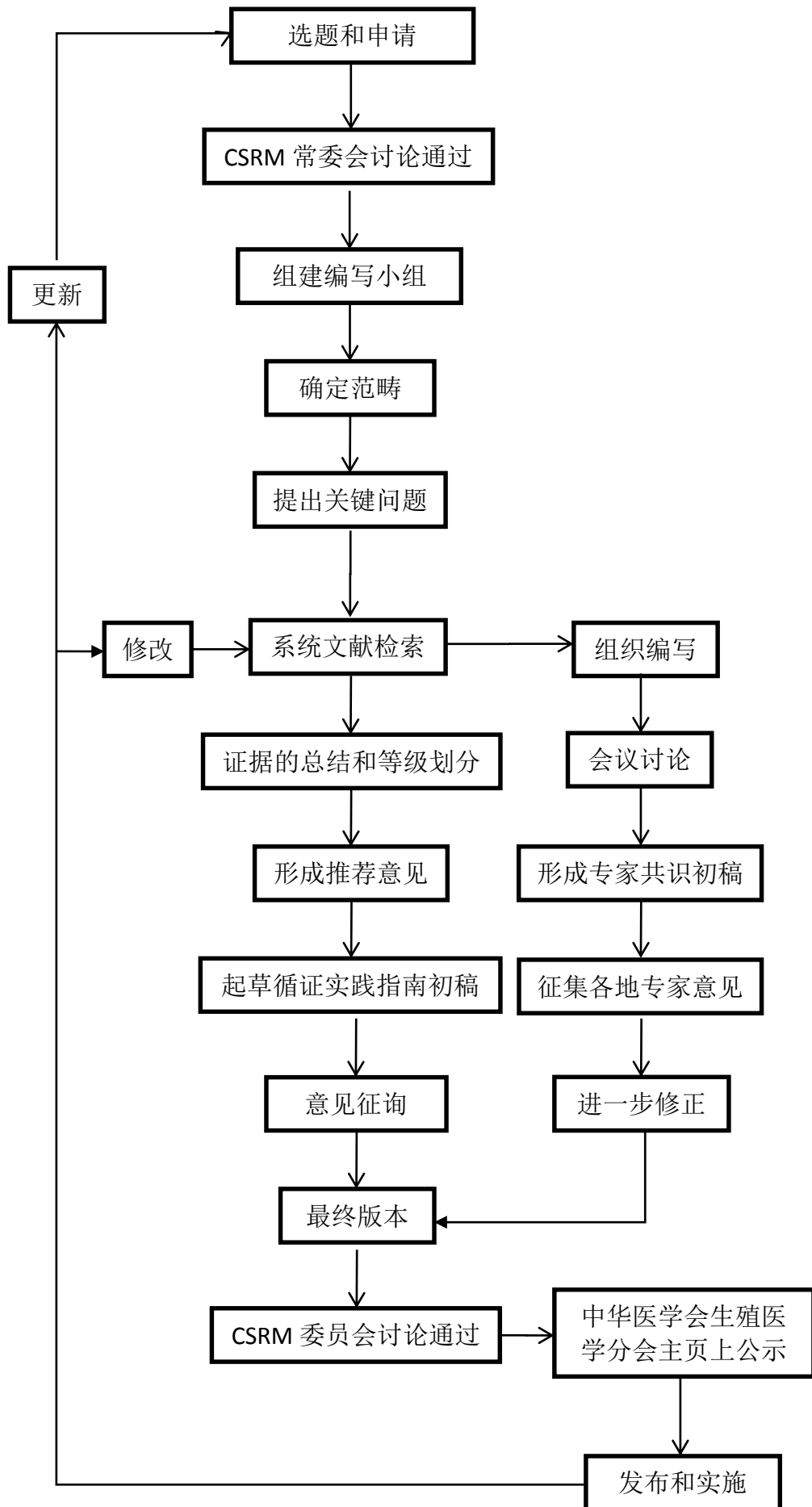
指南共识只有实施下去，对临床操作才有真正的影响。CSRM 可以对指南共识的制定和实施负责，但不能直接对每个生殖中心的具体实施情况进行管理。因此，指南共识发布和实施一段时间后，可以通过一些指标进行评价其影响和效果。

九. CSRM 指南共识的更新

指南共识的更新关注于对建议的实质性的更改，而不是编辑修改文档。可能因以下原因需要更新：

- 可用性的干预
- 分析的结果
- 现存利益的证据和在干预措施的不足
- 可用性资源

指南需要不断的更新，为了制定新版本的指南，申请程序仍需要按照本手册重新进行。



参考文献

1. 中华医学杂志社。《关于中华医学会系列杂志指南共识类文章撰写与发表的推荐规范》。
2. 赵萌。医学专家们,您对临床实践指南概念及国际临床指南知多少?中国组织工程研究Chinese Journal of Tissue Engineering Research, August 27,2015 Vol. 19, No.35.
3. Field MJ, Lohr KN. Guidelines for clinical practice: from development to use [M]. Washington DC: National Academy Press,1992:27.
4. 卫茂玲, 刘鸣. 中国循证医学杂志, 2013, 13(8): 927~932.
5. Atkins D, Eccles M, Flottorp S, et al. Systems for grading the quality of evidence and the strength of recommendations I : critical appraisal of existing approaches The GRADE Working Group. [J]. BMC Health Serv Res, 2004, 4(1):38.

附录 A：申请表

1. 申请人

2. CSRM 的哪个学组参与

3. 提出何种临床问题

4. 涉及相关的临床问题（数量，经费和患者的影响）

5. 拟定的指南共识能解决何种问题

6. 拟定指导方针在发展和实施中将会得到何种预期收益

7. 提出的问题领域是否存在现有指南共识（查询：www.g-i-n.net，www.guideline.gov，www.asrm.org，www.nice.org.uk，www.rcog.org.uk）

8. 请提供证明提案的证据的可信性（查询 Cochrane 系统评价数据库）

签名日期

附录 B：保密表格

作为一个 CSRM 指南的作者，你已经或可能暴露于某些机密或专有信息，材料或资料。在任何情况下，本资料须严格保密，不得在任何情况泄露，这对本资料的完成和最终结果的完整性很重要。

编写小组成员或编辑者的姓名
地址，电话和邮箱
CSRM 组织中所任职务

我不会透露任何与指南编写团队工作相关的机密和/或专有信息、材料或数据给任何第三方，并保证信息的严格保密。

我将把我所持有的任何机密和/或专有信息、材料或数据保存在安全的地方，以防止意外泄露。

除参与 CSRM 指南共识编写工作外，我将不会使用任何机密信息和/或专有信息、材料或数据。

签名日期

附录 C：范畴清单

1. 整体目标
临床指南共识被定义为“在特定的临床环境下，针对特定的临床疾病，系统制定出帮助医务人员和患者对患者本身的健康问题做出恰当的认识和处理提供指导性意见”
2. 目标使用者
哪些临床专业？患者？专业人员？政策制定者？
包括：
不包括：
3. 患者人群
临床实体，诊断类型？病程/等级？年龄限制？
包括：
不包括：
4. 医疗保险设置
二级或三级护理？立法？获得照顾、赔偿？
包括：
不包括：
5. 预防性干预是否应当编入
生活方式？治疗取消？筛选？
6. 应包括哪些诊断干预
哪些诊断干预措施可行？在当前的实践中的变量？新的或者实际上的？
包括：
不包括：
7. 应包括哪种治疗干预
哪些治疗干预措施可行？在当前的实践中的变量？新的或者实际上的？
包括：
不包括：
8. 需要说明哪种结果
哪些结果是常用的报告？哪些结果能够反应实际变化？新的或实际上的？病人更倾向于哪种结果衡量？并发症？成本？
包括：

不包括:
9. 推荐患者人群更易发生什么? 是否已经包括?
诊断干预? 治疗干预? 结果衡量? 并发症 (如双胞胎, OHSS 的风险)? 成本 (如双胞胎, 额外的治疗周期)?
额外包括:
10. 哪些方法将会用来衡量患者喜好?
文献搜索? 咨询? (方式: 群体访谈, 个人访谈, 调查)
包括:
不包括:
11. 与其他文件的关系
其他国家指南? CSRM 的伦理与法律任务强制? 其它法律?
12. 遗漏的内容?
信息提供? 实习机构?
13. 在一定时间内是否能够完成拟定方案?
<u>如果不行</u> : 分割方案? 更多的限制/排除?

附录 D1：质量评估

清单：系统综述

研究鉴定：

步骤 1：评估是否合适	判断
文章是系统的综述还是 meta 分析？	<input type="checkbox"/> 是 =>继续 <input type="checkbox"/> 否 =>放弃
文章与关键问题是否相关？ 用 PICO 分析	<input type="checkbox"/> 是 =>继续去清单 <input type="checkbox"/> 否 =>放弃

步骤 2：方法学评估	判断
综述是否描述了适合的清楚的定义问题？	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否
综述是否包括了你认为与拟定方针相关的研究种类？	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否
综述的审查方法是否描述充分，方法是否合适？	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否
文献是否能够完全定义相关研究？	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否
所包含研究的科学质量是否被评估和报道过？	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否
综述的总体方法学评估（移至证据表）	
这篇综述的方法学质量的总体评估结果是什么？	<input type="checkbox"/> 高质量（1++） <input type="checkbox"/> 可以接受（1+） <input type="checkbox"/> 不能接受-拒绝（1-） ¹
评论： 例如：该文是高质量的与关键问题相关的系统综述。	

附加信息：

本清单的目的是考量解答指南关键问题的系统综述的适用性。这一评估包含 2 个方面：首先，综述描述的问题是否适于解答关键问题，其次，综述的方法学是否足够充分以得到有效的结论。

¹ 如果综述不符合某些标准，它仍然可以用来参考，但不能完全只依赖于它本身去解答综述问题。

附录 D2：质量评估

清单：随机控制试验

研究鉴定：

步骤 1：评估是否合适	判断
文章与关键问题是否相关？ 用 PICO 分析	<input type="checkbox"/> 是 =>继续去清单 <input type="checkbox"/> 否 =>放弃

步骤 2：方法学评估	判断
A. 选择偏差（各组的系统差异）	
随机抽样法是为了分配参与者的治疗组（这可以平衡各组间的干扰因素）	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
有充足的隐蔽性分配（这样调查者/参与者将不能影响参与或分配治疗）	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
基线可比性，包括所有的混杂和预后因素	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
基于以上回答，是否存在选择偏差，如果这样，将有哪些影响？	<input type="checkbox"/> 是 <input type="checkbox"/> 否
B. 性能偏差（除了调查中的干预，在护理提供方面的群体间系统差异）	
除了干预外，调查组提供相同护理	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
“盲”分参与者的护理照顾	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
“盲”分参与者的个人管理照顾	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
基于以上回答，是否存在性能偏差，如果这样，将有哪些影响？	<input type="checkbox"/> 是 <input type="checkbox"/> 否
C. 损耗偏差（比较组间病人流失差异）	
所有组随访时间相同（或者调整至随访期间所能容忍的差异范围）	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
各组间的治疗完成比较（各组有多少参与者没有完成治疗）	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
数据结果的可用性比较（每组有多少参与者没有结果数据可用）	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
基于以上回答，是否存在损耗偏差，如果这样，	<input type="checkbox"/> 是

将有哪些影响?	<input type="checkbox"/> 否
D. 检测偏差（结果在检查、确认和诊断中的偏差）	
调查组的随访时间相似	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
研究使用精确的结果定义	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
有效可靠地方法来确定结果	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
研究者”盲”对参与者的调查和其他重要混杂因素和影响预后的因素	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
基于以上回答，是否存在检测偏差，如果这样，将有哪些影响?	<input type="checkbox"/> 是 <input type="checkbox"/> 否
整体系统学评估（移至证据表）	
存在何种偏差	<input type="checkbox"/> 选择偏差 <input type="checkbox"/> 性能偏差 <input type="checkbox"/> 损耗偏差 <input type="checkbox"/> 检测偏差
这篇综述的方法学质量的总体评估结果是什么?	<input type="checkbox"/> 高质量（1++） <input type="checkbox"/> 可以接受（1+） <input type="checkbox"/> 不能接受-拒绝（1-）
评论： 例如：该文来自多个随机试验，是高质量 RCT 研究。	

附加信息：

这个清单的目的是考量 RCT 对于回答指南关键问题的适用性。这样的评价有 2 个方面：首先，RCT 描述的问题是否适于解答关键问题，其次，RCT 的方法学是否足够充分以得到有效的结论。

附录 D3：质量评估

清单：队列研究

研究鉴定：

步骤 1：评估是否合适	判断
文章与关键问题是否相关？ 用 PICO 分析	<input type="checkbox"/> 是 =>继续去清单 <input type="checkbox"/> 否 =>放弃

步骤 2：方法学评估	判断
A. 选择偏差（各组的系统差异）	
治疗组的分配方法与潜在的混杂因素无关（即参与者分配给治疗组的原因是不希望影响结果的）	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
在设计和分析中尝试平衡对照组的潜在混杂因素	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
基线可比性，包括所有的混杂和预后因素	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
基于以上回答，是否存在选择偏差，如果这样，将有哪些影响？	<input type="checkbox"/> 是 <input type="checkbox"/> 否
B. 性能偏差（除了调查中的干预，在护理提供方面的群体间系统差异）	
除了干预外，调查组提供相同护理	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
“盲”分参与者的护理照顾	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
“盲”分参与者的个人管理照顾	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
基于以上回答，是否存在性能偏差，如果这样，将有哪些影响？	<input type="checkbox"/> 是 <input type="checkbox"/> 否
C. 损耗偏差（比较组间病人流失差异）	
所有组随访时间相同（或者调整至随访期间所能容忍的差异范围）	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
各组间的治疗完成比较（各组有多少参与者没有完成治疗）	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
数据结果的可用性比较（每组有多少参与者没有结果数据可用）	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
基于以上回答，是否存在损耗偏差，如果这样，将有哪些影响？	<input type="checkbox"/> 是 <input type="checkbox"/> 否

D. 检测偏差（结果在检查、确认和诊断中的偏差）	
调查组的随访时间相似	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
研究使用精确的结果定义	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
有效可靠地方法来确定结果	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
研究者”盲”对参与者的调查和其他重要混杂因素和影响预后的因素	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
基于以上回答，是否存在检测偏差，如果这样，将有哪些影响？	<input type="checkbox"/> 是 <input type="checkbox"/> 否
整体系统学评估（移至证据表）	
存在何种偏差	<input type="checkbox"/> 选择偏差 <input type="checkbox"/> 性能偏差 <input type="checkbox"/> 损耗偏差 <input type="checkbox"/> 检测偏差
这篇综述的方法学质量的总体评估结果是什么？	<input type="checkbox"/> 高质量（1++） <input type="checkbox"/> 可以接受（1+） <input type="checkbox"/> 不能接受-拒绝（1-）
评论： 	

附录 D4：质量评估

清单：病例对照研究

研究鉴定：

步骤 1：评估是否合适	判断
文章与关键问题是否相关？ 用 PICO 分析	<input type="checkbox"/> 是 =>继续去清单 <input type="checkbox"/> 否 =>放弃

步骤 2：方法学评估	判断
研究描述了一个适宜的清晰的焦点问题	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否
A. 参与者的选择	
病例和对照来自可比人群	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
病例和对照使用同种排除标准	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
每组的参与比例是多少（病例和对照）	
参与者与非参与者比较以估计相似性和差异	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
病例清晰定义，与对照不同	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
清楚的评估对照不是病例	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
B. 评估	
采取措施防止初级暴露知识影响案件认定	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
采用标准、有效和可信的方法评估暴露状态	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
C. 混杂因素	
鉴定主要潜在混杂因素，并在设计和分析中考虑进去	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
D. 数据分析	
有无置信区间	<input type="checkbox"/> 是 <input type="checkbox"/> 不清楚 <input type="checkbox"/> 否 <input type="checkbox"/> N/A
整体系统学评估（移至证据表）	

存在何种偏差	<input type="checkbox"/> 选择患者偏差 <input type="checkbox"/> 评估偏差 <input type="checkbox"/> 混杂因素偏差
这篇综述的方法学质量的总体评估结果是什么？	<input type="checkbox"/> 高质量（1++） <input type="checkbox"/> 可以接受（1+） <input type="checkbox"/> 不能接受-拒绝（1-）
评论：	

附录 E1：干预研究的证据表

参考书目	研究类型	研究质量 经费+竞争利益	患者编号 患者特点+ 组相似性	干预措施 (+比较) 包括:研究 时间/随访	结果评估 包括:危害 /不良事件	效应大 小	作者结 论	评论
Shaker, et al. FertilSte ril 1996;65: 992-6	RCT	<input type="checkbox"/> 选择偏差 <input type="checkbox"/> 性能偏差 <input type="checkbox"/> 损耗偏差 <input type="checkbox"/> 检测偏差 <input type="checkbox"/> 高质量 (1++) <input type="checkbox"/> 可以接受 (1+) <input type="checkbox"/> 不能接受-拒绝 (1-)	N=22 IVF 病人 OHSS 风险定义为 E2>3,540pg/ml 34 岁 上下浮动 4.4 岁	取卵日和术 后 5 天输注 白蛋白+鲜胚 移植	严重 / 中度 OHSS (1) 临床妊娠/女 性 (2)	OR (95% CI) 5.3 (0.5-56.2) 0.06 (0.01-1.2)	胚胎冷冻 与输注白 蛋白预防 OHSS 同 样有效, 并获得更 高临床妊 娠率	NA

填表说明：

当没有可添加因素，包括：

“不适用 (NA)” 当一个项目是不通知，或

“未描述 (ND)” 当一个项目必须被告知，但在出版物中没有给出信息

参考书目 (+PMID) 作者，杂志，年，卷和页 +PMID

研究类型 Meta 分析/综述 (SR)，随机对照试验 (RCT)，非随机队列研究 (CS)，
病例报告 (CASE)，意见书 (OPINION)，其他 (请注明) …

研究质量 从质量评估中添加结果 (1++,1+,1-,2++,2+, 2-,3,4)

经费和竞争利益 标明经费来源。CI：写明“开始”或者“没有开始”并说明

评论 质量评估中的重要评论，或者出版类型、经费或竞争利益等的评论

患者编号 患者总数，每组患者数

患者特点+组相似性 描述相关基本特点 (年龄、疾病状况、纳入/排除标准等)

干预措施 (+比较) 指定每个组的干预 (治疗/程序 (剂量，方案，长度，..)/安慰剂/替代治
疗/期待的管理)

研究时间/随访 将参与研究患者纳入到一个指定的终点 (如移植) 或数据采集终端，提
到有关结果

结果评估 (主要/次要) 所有的测量衡量 (阳性和阴性)：例如 OHSS 发生率，植入率，妊娠，
生活质量，满意度，…

如作者指定，需区分主要 (1) 和次要 (2) 结果

效应大小 绝对风险降低，相对风险，需要治疗/伤害的数字，或比值比与置信区间；
清楚描述使用哪种结果衡量。如可能，请添加 p 值。

危害/不良事件 像文献中一样定义和描述每组观察到的伤害。精确的平均值，百分比，
p 值。

作者结论 陈述作者结论

评论 这项研究的其他/解释或缺陷特征，审稿人提出的其他计算方法 (NNT，
RR，或者，CI，..)

附录 E2：用于诊断研究的证据表

参考书目	研究类型	研究质量 经费+竞争 利益	患者编号 患者特点+组 相似性	诊断试验评价 参考标准试验 包括：时间间隔 和治疗	患病率	精确度 (Se, Sp, PPV, NPV, LR+, LR-)	再现性	作者 结论	评论
Lee, et al. Hum Reprod 2008;23: 1	CS	<input type="checkbox"/> 高质量 (1++) <input type="checkbox"/> 可以接受 (1+) <input type="checkbox"/> 不能接受-拒绝 (1-)	N=262 一个周期控制性卵巢刺激排除：超长/短或拮抗剂，赠卵	基础血清 AMH 参考：Navot 鉴定的中度或重度 OHSS	8%	敏感性 91%， 特异性 81%， PPV30% NPV 99%	NA		

填表说明：

当没有可添加因素，包括：

“不适用 (NA)” 当一个项目是不通知，或

“未描述 (ND)” 当一个项目必须被告知，但在出版物中没有给出信息

参考书目 (+PMID) 作者，杂志，年，卷和页 +PMID

研究类型 Meta 分析/综述 (SR)，随机对照试验 (RCT)，非随机队列研究 (CS)，病例报告 (CASE)，意见书 (OPINION)，其他 (请注明) …

研究质量 从质量评估中添加结果 (1++,1+,1-,2++,2+, 2-,3,4)

经费和竞争利益 标明经费来源。CI：写明“开始”或者“没有开始”并说明

评论 质量评估中的重要评论，或者出版类型、经费或竞争利益等的评论

设置 多中心，国家，医疗保健机构设置

患者编号 患者总数，每组患者数

患者特点+组相似性 描述相关基本特点 (年龄、疾病状况、纳入/排除标准等)

干预措施 (+比较) 指定每个组的干预 (治疗/程序 (剂量，方案，长度，..)/安慰剂/替代治疗/期待的管理)

诊断试验评价 描述评估试验 (什么，由谁和如何，时间，..)，切断，盲对临床信息和/或指数测试结果

参考标准试验 描述评估试验 (什么，由谁和如何，时间，..)，切断，盲对临床信息和/或指数测试结果

时间间隔和治疗 指定测试之间的时间间隔和治疗 (如果有的话)

患病率 疾病患病率

精确度 给出所有含置信区间的可用图表 (如果可用)：

灵敏度 (Se)，特异性 (SP)，阳性预测值 (PPV)，阴性预测值 (NPV)，似然比 (LR, LR-)，ROC 曲线下面积

再现性 给出所有含置信区间的可用图表 (如果可用)：

定量试验 (重复次数，范围，值，协议，相关性)

定性试验 (可靠性，相关系数，..)

危害/不良事件 像文献中一样定义和描述每组观察到的伤害。精确的平均值，百分比，p 值。

作者结论 陈述作者结论

评论 这项研究的其他/解释或缺陷特征，其他计算方法 (NNT, RR, 或者, CI, ..)

