

ICU 谵妄评估工具诊断准确性系统评价再评价

崔念奇¹, 郑琼², 张玉萍¹, 金静芬^{1,3,4}

Diagnostic accuracy of delirium assessment tools in intensive care unit: an umbrella review of systematic reviews and meta-analysis Cui

Nianqi, Zheng Qiong, Zhang Yuping, Jin Jingfen

摘要:目的 对 ICU 谵妄评估工具诊断准确性系统评价进行再评价。方法 2022 年 4 月,计算机检索 7 个英文和 3 个中文数据库中 与 ICU 谵妄评估工具诊断准确性相关的系统评价,由 2 名研究人员独立进行文献筛选,资料提取,方法学及证据质量评价。通过 GRADE pro GDT(Guideline Development Tool)在线工具进行证据质量分级。使用重叠百分比,重叠区域及修正重叠区域对纳入系统评价中的原始研究重复情况进行分析。结果 共纳入 5 篇系统评价,3 篇为高质量,2 篇为中等质量。整体证据质量级别为低和极低。ICU 意识模糊评估法合并敏感度为 0.77~0.84,合并特异度为 0.95~0.97。纳入研究修正重叠区域为 0.24。结论 ICU 意识模糊评估法可能是 ICU 患者谵妄评估的最佳工具,但证据质量级别低。亟待构建诊断准确性系统评价质量评价工具,未来指南改编可应用系统评价再评价为推荐意见的形成提供证据来源。

关键词:谵妄;重症监护室;量表;评估工具;敏感度;特异度;系统评价再评价

中图分类号:R473.5 **文献标识码:**B **DOI:**10.3870/j.issn.1001-4152.2022.15.032

谵妄是一种急性发作和病情反复波动的疾病,其特征是意识和注意障碍,伴有认知功能的改变、睡眠-觉醒周期紊乱和精神运动性兴奋或抑制^[1]。一项针对 ICU 成人患者的系统评价报道 ICU 谵妄发生率为 31%^[2]。谵妄可对 ICU 患者造成诸多的负性影响,如增加 ICU 住院时间、持续认知功能障碍及病死率。谵妄的持续监测和早期筛查是谵妄治疗的重要基础,ICU 护士在谵妄评估及筛查中扮演关键角色,亟需提供兼具高灵敏度和特异度的谵妄评估工具^[3]。随着 ICU 谵妄评估工具诊断准确性试验数量的增加,研究者对不同的评估工具进行了诊断准确性系统评价及 Meta 分析,但其证据级别尚不清楚,研究结论不统一。美国医学研究所(Institute of Medicine, IOM)将临床实践指南定义为:通过系统评价生成的证据以及对各种备选干预方式进行利弊评价之后提出的最优指导意见^[4]。故所有推荐意见的证据都应源于系统评价。然而针对同一临床问题可能会存在诸多系统评价,决策者根据海量的系统评价证据进行决策时可能会面临障碍。系统评价再评价(umbrella reviews)通过对特定研究问题可获得的系统评价和 Meta 分析进行全面而系统的评价,形成特定研究问题或现有研究结果的全面概况,并提供了解证据体强度和潜在偏倚风险的可能性^[5]。就 ICU 患者谵妄评估而言,有别于诊断准确性试验的 Meta 分析往往仅能提供单一评估工具诊断的准确性,系统评价再评价可以综合不同谵妄评估工具的诊断准确性,并比较同

一评估工具不同研究结果之间的差别,明确不同系统评价及 Meta 分析证据级别及可能的偏倚风险。因此,本研究基于 JBI 系统评价再评价方法学^[5],对现有适合 ICU 患者谵妄评估工具诊断准确性系统评价及 Meta 分析进行全面评价,明确证据质量及潜在偏倚风险,为护理人员进行临床决策提供参考。研究计划书已于 PROSPERO 平台注册(ID:CRD42021242586)并公开发表^[6]。本研究基于系统评价再评价优先报告条目清单(Preferred Reporting Items for Overviews of Systematic Reviews including harms checklist, PRIO-harms)进行报告^[7-8]。

1 资料与方法

1.1 文献纳入与排除标准 纳入标准:①研究类型,诊断准确性试验系统评价及 Meta 分析。②研究对象,ICU 成人患者(年龄≥18 岁),病种不限。③诊断标准,采用公认的美国精神疾病诊断与统计手册(DSM)^[1]或 WHO 国际疾病分类(ICD)^[9]作为金标准,具体版本不限。以 DSM-V 为例,谵妄诊断标准包括意识障碍,伴有注意力的集中、保持或变换目标能力降低;认知的改变或出现知觉障碍,且其认知障碍不能由一种原本存在或发展的神经认知障碍或昏迷所解释;在短时间内发生并在 1 d 内有所波动;从病史、实验室检查或体检中有迹象表明是一般躯体情况的直接生理性后果。④结局指标,合并敏感度和合并特异度。某一谵妄评估工具的合并敏感度=所有真阳性人数/(所有真阳性人数+所有假阴性人数),合并特异度=所有真阴性人数/(所有真阴性人数+所有假阳性人数)。排除标准:研究对象仅纳入 ICU 临终患者,非中英文发表,无法获取全文的文献。

1.2 文献检索策略 2022 年 4 月,计算机检索 CD-SR(Cochrane Database of Systematic Reviews)、JBI、Medline、Embase、CINAHL、Web of Science、PsycINFO、中国知网、维普全文数据库以及中国生物

作者单位:1. 浙江大学医学院附属第二医院护理部(浙江 杭州, 310009);2. 浙江大学医学院护理系;3. 浙江大学医学院附属第二医院长兴院区;4. 浙江省严重创伤与烧伤诊治重点实验室

崔念奇:男,博士在读,学生

通信作者:金静芬, zjzkh1@zju.edu.cn

科研项目:中华护理学会 2019 年度科研课题(ZHKY201913);浙江大学博士研究生学术新星培养计划项目(202059)

收稿:2022-04-04;修回:2022-05-20

医学文献数据库。灰色文献检索 Google Scholar,同时手工检索纳入文献的参考文献列表。检索时限为 2016 年 1 月至 2022 年 4 月,研究类型限制为系统评价和/或 Meta 分析。中文检索词为谵妄、神经认知障碍、苏醒谵妄、意识模糊、急性脑功能障碍、脑代谢疾病;重症监护病房、危重患者医疗、危重症护理、危重症监护结局、重症监护、重症医学科、加强监护病房、ICU、EICU、SICU、CCU;系统评价、Meta 分析、荟萃分析、元分析、综述。英文检索词为 delirium, emergence delirium, neurocognitive disorders, brain syndrome, cerebral insufficiency, delirious *, attention, cognitive, consciousness, perception, neurocognitive; intensive care units, critical care, critical care nursing, intensive care, acute care, cardiac, coronary, heart, burn, respiratory, surgical, surger *, step-down, speciali # ed, highdependency, weaning, postoperati *, postsurg *, ICU *, SICU *, CCU *, EICU *, HDU *, SDU *, EDS DU; Meta-Analysis, Systematic Review,具体检索策略见研究计划书^[6]。

1.3 文献筛选与资料提取 由 2 名研究人员按照标准独立筛选文献,如遇分歧由第 3 名研究人员协助解决。文献筛选参照 PRISMA2020 文献筛选流程^[10]。双人进行文献资料提取,如遇分歧由第 3 名研究人员协助解决。提取内容包括第一作者、诊断准确性试验数量、诊断金标准、谵妄评估工具、合并敏感度、合并特异度、文献质量评价工具。

1.4 质量评价工具及证据分级 由于目前暂无诊断准确性试验系统评价特异性质量评价工具,本研究采用 JBI 系统评价方法学质量评价工具进行质量评价^[11-12]。证据质量的评价方法采用 GRADE^[13],对偏倚风险、间接性、不一致性、不精确性及发表偏倚进行评价。偏倚风险评价采用原系统评价诊断准确性研究质量评价结果,若原系统评价未对偏倚风险进行评估,本研究采用 Cochrane 协作网推荐的诊断准确性研究质量评价工具 QUADAS-2^[14-15]。间接性主要根据纳入系统评价与系统评价再评价拟回答的临床问题之间的匹配程度进行评价。不一致性主要根据合并敏感度、合并特异度可信区间重叠程度, I^2 和 P 值进行评价。不精确性通过系统评价理论所需样本量,合并敏感度、合并特异度可信区间阈值下限与临床可接受阈值下限比较。本研究中通过查表法估算理论所需样本量^[16],临床可接受阈值下限设定为 0.70。

1.5 资料分析方法 使用重叠百分比,重叠区域(Covered Area, CA)及修正重叠区域(Corrected Covered Area, CCA)对纳入系统评价中的原始研究重复情况进行分析^[17]。CCA 在 0~5 表示低度重叠,6~10 中度重叠,11~15 高度重叠,>15 极高度重叠。GRADE 证据质量评价通过 GRADE pro GDT (Guideline Development Tool)进行分级并制作证据

概要表。若纳入系统评价未对发表偏倚进行评估,本研究通过将原始研究数据录入 Stata14.0 制作 Deeks 漏斗图以评估发表偏倚。

2 结果

2.1 文献筛选流程及纳入系统评价特征 初检获得 2 961 篇文献,使用软件去除重复文献,阅读标题、摘要和全文排除主题不符、非 ICU、非系统评价或 Meta 分析文献,最终纳入 5 篇系统评价^[18-22],涉及 ICU 意识模糊评估法(Confusion Assessment Method for the Intensive Care Unit, CAM-ICU)、ICU 谵妄筛查量表(Intensive Care Delirium Screening Checklist, ICDSC)、意识模糊量表(Neelon and Champagne Confusion Scale, NEECHAM)、护理谵妄筛查量表(Nursing Delirium Screening Scale, Nu-DESC)、谵妄识别评分(Delirium Detection Score, DDS)5 种谵妄评估工具,其中一篇系统评价^[21]还涉及 4 项谵妄快速诊断方案(the 4 'A's Test, 4AT),但由于其金标准不符合纳入标准,未提取相关内容。3 篇^[18,20,22]使用 QUADAS-2,1 篇^[24]使用 JBI 诊断性研究质量评价工具,1 篇^[21]使用改良 Newcastle-Ottawa 量表评价方法学质量。纳入系统评价特征见表 1。

2.2 纳入系统评价方法学质量评价 3 篇^[18-19,22]系统评价质量等级为高,即所有条目评价结果均为“是”;2 篇^[20-21]系统评价质量等级为中,其中 Jeong 等^[20]发表的系统评价未评估发表偏倚,Patel 等^[21]发表的系统评价未采用合适的质量评价标准对原始研究进行方法学质量评价。

2.3 纳入系统评价重复研究情况 重叠百分比为 $25/51 \times 100\% = 49.02\%$,CA 为 $80/(51 \times 5) = 0.31$,CCA 为 $(80-51)/(51 \times 5 - 51) = 0.14$ 。

2.4 证据质量评价 对纳入系统评价中不同的谵妄评估工具合并敏感度及合并特异度进行质量评价,结果显示谵妄评估工具中等质量证据占比 5%,低质量证据占比 75%,极低质量证据占比 20%。证据质量降级主要集中在偏倚风险,不一致性和不精确性。仅 Jeong 等^[20]的系统评价未评估纳入研究的发表偏倚,本研究通过制作 Deeks 漏斗图评估发表偏倚,结果显示 P 值为 0.19,提示漏斗图对称,存在发表偏倚可能性较低,不考虑降级。

3 讨论

3.1 ICU 患者谵妄评估工具证据质量 本研究纳入的 5 篇系统评价包含多个适用于 ICU 患者的谵妄评估工具,均显示具有良好的敏感度和特异度,但证据质量普遍为低或极低。重复研究结果显示,本研究纳入系统评价存在低度重叠现象,研究结果效应因重叠被过度放大的可能性较低。基于现有证据,CAM-ICU 可能是最适合于 ICU 患者谵妄评估的工具。在 Chen 等^[18]的研究中,CAM-ICU 相较于 ICDSC,发现非谵妄患者及内科 ICU 和机械通气谵妄患者表现更

佳($P<0.05$)。Ho 等^[19]的研究同样支持 CAM-ICU 是现有最佳 ICU 患者谵妄评估工具,但是该研究未对不同的谵妄评估工具诊断准确性差别进行显著性检验。Patel 等^[21]发现,CAM-ICU 对神经重症患者谵妄评估同样具有较好的适用性。本研究表明,相较于 DDS,CAM-ICU 在老年重症患者中具有良好的敏感度和特异度。与 CAM-ICU 相比,Nu-DESC 合并敏感度和特异度的表现不佳,证据质量等级为极低到低质量。证据质量对指南推荐意见强度具有重要影响,但并不是唯一重要因素,利益相关人群的价值观与偏好同样是影响指南推荐意见的重要因素。Jeong 等^[20]推荐使用 Nu-DESC 作为谵妄评估工具的原因是:Nu-DESC 兼具诊断准确性和临床适用性,是“实践者友好型”谵妄评估工具,即 Nu-DESC 更加符合临床护士价值观与偏好。相关研究报告,ICU 护士在使

用 CAM-ICU 对患者进行谵妄评估时存在障碍,认为 CAM-ICU 使用过程存在困难^[23-24]。Tsuruta 等^[25]推荐将多种评估工具相结合进行谵妄筛查,然而现有谵妄筛查障碍因素之一就是增加护理人员工作负担^[3]。鉴于 CAM-ICU 在谵妄评估中的诊断准确性结果,如何增加临床护士应用 CAM-ICU 对谵妄筛查的执行力是未来的研究重点,同时政策制定者在向实践者推荐谵妄评估工具时,应充分考虑不同评估工具证据质量及利益相关人群价值观。美国重症医学会 2018 年发布的 PADIS 指南^[8]在谵妄评估工具的推荐意见给予良好实践声明,未明确具体的证据质量及推荐级别,主要因为谵妄筛查效果与患者终点结局之间的相关性存在证据缺口,因此未来研究可针对这一证据缺口开展相关研究。

表 1 纳入系统评价特征

纳入研究	纳入文献数量		评估工具	JBI 等级	结果	降级证据质量的因素					GRADE 等级
	(个)/样本量 (例)	金标准				偏倚 风险	间接性	不一致性	不精确性	发表 偏倚	
Chen 等 ^[18]	30/4007	DSM/ICD	CAM-ICU	高	合并敏感度 0.84(95%CI:0.77~0.88)	严重 ^a	不严重	严重 ^b	不严重	无	低
					合并特异度 0.95(95%CI:0.91~0.97)	严重 ^a	不严重	严重 ^c	不严重	无	低
	12/1247	DSM/ICD	ICDSC		合并敏感度 0.83(95%CI:0.74~0.90)	严重 ^d	不严重	严重 ^e	不严重	无	低
Ho 等 ^[19]					合并特异度 0.87(95%CI:0.78~0.93)	严重 ^d	不严重	严重 ^f	不严重	无	低
	23/3584	DSM	CAM-ICU	高	合并敏感度 0.85(95%CI:0.77~0.91)	严重 ^g	不严重	严重 ^h	不严重	无	低
					合并特异度 0.95(95%CI:0.90~0.97)	严重 ^g	不严重	严重 ^h	不严重	无	低
	8/1206	DSM	ICDSC		合并敏感度 0.87(95%CI:0.70~0.95)	严重 ⁱ	不严重	严重 ^j	不严重	无	低
					合并特异度 0.91(95%CI:0.85~0.95)	严重 ⁱ	不严重	严重 ^j	不严重	无	低
Jeong 等 ^[20]	1/105	DSM	NEECHAM		敏感度 0.97(95%CI:0.85~1.00)	不严重	不严重	不严重	非常严重 ^k	无	低
					特异度 0.83(95%CI:0.73~0.90)	不严重	不严重	不严重	非常严重 ^k	无	低
	11/2245	DSM	Nu-DESC	中	合并敏感度 0.69(95%CI:0.55~0.80)	不严重	严重 ^l	严重 ^m	严重 ⁿ	无	极低
Patel 等 ^[21]					合并特异度 0.89(95%CI:0.83~0.94)	不严重	严重 ^l	严重 ^m	不严重	无	低
	2/1064	DSM	CAM-ICU	中	合并敏感度 0.77(95%CI:0.72~0.82)	严重 ^o	不严重	严重 ^p	不严重	无	低
					合并特异度 0.97(95%CI:0.96~0.98)	严重 ^o	不严重	严重 ^q	不严重	无	低
	1/61	DSM	ICDSC		敏感度 0.64(95%CI:0.49~0.77)	严重 ^r	不严重	不严重	非常严重 ^s	无	极低
					特异度 0.79(95%CI:0.63~0.89)	严重 ^r	不严重	不严重	非常严重 ^s	无	极低
van Velthuis 等 ^[22]	2/1225	DSM	DDS	高	合并敏感度 0.54(95%CI:0.46~0.62)	严重 ^t	不严重	严重 ^u	非常严重 ^v	无	极低
					合并特异度 0.77(95%CI:0.74~0.79)	严重 ^t	不严重	严重 ^w	不严重	无	低
	2/221	DSM	CAM-ICU		合并敏感度 0.84(95%CI:0.73~0.91)	不严重	不严重	不严重	非常严重 ^x	无	低
					合并特异度 0.95(95%CI:0.91~0.98)	不严重	不严重	不严重	严重 ^y	无	中

注:a. QUADAS-2 评估结果中,1 项研究病例选择条目风险不确定;9 项研究待评价试验条目风险不确定;10 项研究金标准条目风险不确定;5 项研究病例流程和进展情况条目风险不确定,1 项研究存在高风险。b. 合并敏感度 I^2 为 87.99%($Q=241.53, P<0.001$)。c. 合并特异度 I^2 为 96.64%($Q=863.78, P<0.001$)。d. QUADAS-2 评估结果中,1 项研究病例选择条目风险不确定;4 项研究待评价试验条目风险不确定,1 项研究存在高风险;4 项研究金标准条目风险不确定;6 项研究病例流程和进展情况条目风险不确定,1 项研究存在高风险。e. 合并敏感度 I^2 为 82.70%($Q=63.57, P<0.001$)。f. 合并特异度 I^2 为 87.93%($Q=91.10, P<0.001$)。g. JBI 诊断性研究评估结果中,19 项研究在界值条目风险不确定。h. 合并敏感度及合并特异度 I^2 均大于 50%。i. JBI 诊断性研究评估结果中,6 项研究在界值条目风险不确定。j. 合并敏感度及合并特异度 I^2 均大于 50%。k. 样本量低于理论所需最低样本量,可信区间较宽。l. 6 项原始研究的实施地点为普通病房,包括康复病房,内科病房,外科病房;3 项在复苏室或麻醉后监护室,1 项在重症监护室,1 项在急诊科。m. 未提供 I^2 值,可信区间重叠程度较差。n. 可信区间阈值下限低于 0.70。o. 改良 Newcastle-Ottawa 量表评估结果中,1 项研究得分为 7 分,另一项研究得分为 8 分。p. 合并敏感度 I^2 为 76.8%($Q=4.30, P=0.038$)。q. 合并特异度 I^2 为 96.8%($Q=30.97, P<0.001$)。r. 改良 Newcastle-Ottawa 量表得分为 7 分。s. 样本量低于理论所需最低样本量,可信区间较宽。t. QUADAS-2 评估结果中,条目 4、6、7、8 条目风险不确定。u. 合并特异度 I^2 为 95.6%($Q=22.95, P<0.001$)。v. 可信区间阈值下限低于 0.70 且可信区间较宽。w. 合并特异度 I^2 为 92.90%($Q=14.16, P<0.001$)。x. 样本量低于理论所需最低样本量,可信区间较宽。y. 样本量低于理论所需最低样本量。

3.2 亟待形成诊断准确性试验系统评价质量评价工具 诊断准确性试验系统评价方法学质量评价目前暂无特异性工具,偏倚风险的评估也缺乏明确标准。因 AMSTAR 2 制定严谨,应用广泛且可操作性强,GRADE 工作组推荐使用该工具进行系统评价方法学质量评价。而 AMSTAR 2 适应范围不包括诊断性试验系统评价、网状 Meta 分析、单个病例数据的 Me-

ta 分析、概况性评价和现实主义评价^[26]。GRADE 工作组推荐使用 QUADAS-2 评价工具对诊断准确性试验系统评价偏倚风险进行质量评价,即从病例选择的偏倚风险、待评价试验的偏倚风险、参考试验(金标准)的偏倚风险及病例流程与进展情况的偏倚风险四方面综合评价^[16]。而 QUADAS-2 适用于诊断准确性试验原始研究的偏倚风险评价,因此该工具在诊断

准确性试验系统评价中的实用性有待验证。鉴于目前暂无针对诊断准确性试验系统评价质量评价工具,亟待形成适合该类系统评价的质量评价工具,帮助研究者进行准确的质量分级。

3.3 系统评价再评价为指南改编推荐意见形成提供重要证据来源 临床实践指南对于医疗质量的提升至关重要,但构建一部新指南对于资源的需求是巨大的^[27]。当已存高质量指南时,指南改编被认为是构建新指南的替代方案^[28]。但由于待改编指南的发布时间及指南纳入证据本身存在滞后性的特点,最新的研究证据往往不能及时被指南纳入。系统评价作为指南改编的重要证据来源,在形成推荐意见的过程中扮演重要角色。系统评价再评价通过对特定临床问题或现象的系统评价进行综合,可以为决策者提供更为全面的研究证据概况,从而更有利于帮助决策者进行临床决策。以最经典的干预性系统评价为例,系统评价再评价相较于系统评价在如下方面具有优势:综合同一种疾病或人群不同干预措施的应用效果;综合同一种干预措施不同方法的干预效果;综合同一种干预措施在不同人群中的应用效果;综合同一干预措施对同一或不同人群的负性影响;同一人群的同一干预措施在不同系统评价结果中存在冲突^[29]。鉴于系统评价再评价的上述优势,本研究团队在指南改编的过程中,将系统评价再评价作为形成推荐意见的重要证据来源。

4 小结

谵妄评估工具诊断准确性试验整体证据质量等级较低。在现有证据中,CAM-ICU 可能是 ICU 患者谵妄评估的最佳工具。但管理者并不能依此推荐该工具,应充分结合利益相关人群价值观、本土适用性等方面综合考虑。未来指南改编可通过系统评价再评价为推荐意见的形成提供证据来源,同时亟待构建诊断准确性试验系统评价质量评价工具。本研究局限性:由于不同原始研究数据报告差异性,本研究未再次提取原始研究数据进行 Meta 分析;研究仅纳入 2016 年之后发表的中英文文献,结果可能存在偏倚。

参考文献:

- [1] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition[M]. Arlington, VA, London: American Psychiatric Publishing, Inc, 2013: 596-598.
- [2] Krewulak K D, Stelfox H T, Leigh J P, et al. Incidence and prevalence of delirium subtypes in an adult ICU: a systematic review and meta-analysis[J]. Crit Care Med, 2018, 46(12): 2029-2035.
- [3] Devlin J W, Skrobik Y, Gelinas C, et al. Clinical practice guidelines for the prevention and management of pain, agitation/sedation, delirium, immobility, and sleep disruption in adult patients in the ICU[J]. Crit Care Med, 2018, 46(9): e825-e873.

- [4] Institute of Medicine. Clinical Practice Guidelines We Can Trust[M]. Washington (DC): National Academies Press (US), 2011: 4.
- [5] Aromataris E, Fernandez R, Godfrey C M, et al. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach[J]. Int J Evid Based Healthc, 2015, 13(3): 132-140.
- [6] Cui N, Chen D, Zhang Y, et al. The effect of non-pharmacological interventions on physical restraint reduction in intensive care units: a protocol for an umbrella review of systematic reviews and meta-analysis[J]. Ann Palliat Med, 2021, 10(6): 6892-6899.
- [7] Bougioukas K I, Liakos A, Tsapas A, et al. Preferred reporting items for overviews of systematic reviews including harms checklist: a pilot tool to be used for balanced reporting of benefits and harms[J]. J Clin Epidemiol, 2018, 93: 9-24.
- [8] 卢存存, 杨丰文, 柯立鑫, 等. 系统评价再评价优先报告条目解读[J]. 中国循证儿科杂志, 2018, 13(3): 236-240.
- [9] World Health Organization. WHO releases new International Classification of Diseases (ICD 11) [EB/OL]. (2018-06-18) [2021-07-21]. [https://www.who.int/news/item/18-06-2018-who-releases-new-international-classification-of-diseases-\(icd-11\)](https://www.who.int/news/item/18-06-2018-who-releases-new-international-classification-of-diseases-(icd-11)).
- [10] Page M J, McKenzie J E, Bossuyt P M, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews[J]. BMJ, 2021, 372: n71.
- [11] Aromataris E, Fernandez R, Godfrey C M, et al. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach[J]. Int J Evid Based Healthc, 2015, 13(3): 132-140.
- [12] 顾莺, 张慧文, 周英凤, 等. JBI 循证卫生保健中心关于不同类型研究的质量评价工具——系统评价的方法学质量评价[J]. 护士进修杂志, 2018, 33(8): 701-703.
- [13] Guyatt G H, Oxman A D, Kunz R, et al. What is "quality of evidence" and why is it important to clinicians? [J]. BMJ, 2008, 336(7651): 995-998.
- [14] Whiting P F, Rutjes A W, Westwood M E, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies[J]. Ann Intern Med, 2011, 155(8): 529-536.
- [15] 郭兰, 张永, 曾宪涛. QUADAS-2 在诊断准确性研究的质量评价工具中的应用[J]. 湖北医药学院学报, 2013, 32(3): 201-208.
- [16] 陈耀龙. GRADE 在系统评价和实践指南中的应用[M]. 兰州: 兰州大学出版社, 2017: 129-130.
- [17] Pieper D, Antoine S L, Mathes T, et al. Systematic review finds overlapping reviews were not mentioned in every other overview[J]. J Clin Epidemiol, 2014, 67(4): 368-375.
- [18] Chen T J, Chung Y W, Chang H C R, et al. Diagnostic accuracy of the CAM-ICU and ICDSC in detecting inten-