AAP 临床实践指南制定转向 GRADE 方法

Corinna J. Rea, MD, MPH,[1,2] Sandra P. Spencer, MD[3,4]

[1]Division of General Pediatrics, Boston Children's Hospital, Boston, Massachusetts; [2]Harvard Medical School, Boston, Massachusetts; [3]Section of Emergency Medicine, Children's Hospital Colorado, Aurora, Colorado; and [4]Department of Pediatrics, University of Colorado School of Medicine, Aurora, Colorado

Address correspondence to: Corinna Rea, MD, MPH, 300 Longwood Ave, Boston, MA 02115. Corinna.Rea@childrens.harvard.edu

临床实践指南的重要性

临床实践指南 (CPG) 在系统评价和评估相关文献后，提出了具体的临床推荐意见，为忙碌的临床医师整合了大量证据，旨在提高医疗质量和价值，并规范临床实践。各类组织正为日益广泛的疾病制定 CPG。因此，最终使用者必须信任 CPG 的制定流程，并确信推荐意见是公正且基于证据的。

从历史上看，CPG 主要基于专家意见和临床经验，而非系统的文献评估 [1]。随着时间的推移，CPG 制定流程日趋严谨，美国国家医学院和指南国际网络等权威组织相继发布了关于 CPG 应具备要素的指导意见。目前存在多种证据评价体系，其在证据权衡和推荐意见形成方面的实践各不相同 [2]。

美国儿科学会 (AAP) 独立或合作制定 CPG。与全球趋势一致，AAP 采用了愈发严谨的方法学来制定 CPG，最终转向“推荐分级的评估、制定与评价 (GRADE)”方法 (https://www.aap.org/thegradeapproach)。本篇述评由指南制定委员会负责人撰写，并随 AAP 首次应用 GRADE 制定的《体重增长不佳诊断与管理临床实践指南》同时发布，对这一过渡进行了阐述。

AAP 的 CPG 制定流程演进

2004 年之前，由于缺乏统一的证据评价框架，AAP 根据指南委员会成员之间的共识程度，为 CPG 推荐意见赋予“共识水平”评分。2004 年，AAP 发布了一项政策声明，提出一套新的更标准化的体系来评价相关文献，并为推荐意见赋予强度。该方法学为 AAP 所独有，借鉴了其他体系，采用明确的标准来制定指南。此前 AAP 的所有 CPG 均使用这一体系，其步骤包括：(1) 确定支持拟议推荐意见的证据质量；(2) 评估实施推荐意见时的获益与危害平衡；(3) 赋予推荐意见强度。指南专家组评价每一项研究的质量、各项研究结果的一致性和稳健性以及效应量级。对证据质量进行评级 (A ～ D 级)，设计良好的随机对照试验 (RCT) 权重最高 (A)，小型病例报告或专家意见的权重最低 (D)。另设 X 类，用于无法进行验证性研究但获益或危害明显占优的特殊情况。随后，指南专家组评估遵循指南推荐意见的预期获益、危害、风险和成本。这些因素的权衡结果会影响推荐意见的级别，从强推荐至不推荐 [3]。

尽管 AAP CPG 委员会已使用该体系 20 年，但仍存在改进空间 [4]。该体系为 AAP 所独用，导致难以将其推荐意见和证据评估与其他框架进行比较。此外，亟需一个质量评级一致性更高的、且在方法学中明确纳入健康公平性考量的体系。针对这些不足，加之全球 CPG 制定向标准化发展的趋势，AAP 最终采

American Academy of Pediatrics
DEDICATED TO THE HEALTH OF ALL CHILDREN™
苏州大学附属儿童医院 苏州大学

The article(s) originally appeared in English in the 2026, Volume 157 of Pediatrics, published by the American Academy of Pediatrics (AAP), and are hereby authorized for translation into Chinese by Children's Hospital of Soochow University.

纳了 GRADE 框架。

GRADE 的优势

认识到需要统一的框架，国际指南制定者代表小组于 2004 年提出 GRADE 方法 [5]。越来越多的组织已采纳 GRADE，包括世界卫生组织、美国医师学会和 Cochrane 协作网。通过采用 GRADE，AAP 减少了学习多种分级体系的需求，并使不同 CPG 之间具有更高的一致性 [6]。

尽管 AAP 原有体系与 GRADE 存在许多共同点，但 GRADE 在评估证据确定性（或质量）以及如何从证据过渡到推荐意见方面，提供了更明确的标准。通过将证据质量评价与推荐意见强度进行分离，GRADE 提高了临床医师、家长、照护者和政策制定者对结果的理解。

GRADE 将证据确定性分为高、中、低、极低 4 个等级。证据确定性可根据明确的因素进行降级或升级。导致降级的因素有 5 个（研究设计的局限性、结果不一致性、间接性、不精确性、发表偏倚），导致升级的因素有 3 个（效应量很大、剂量反应关系、混杂因素可能降低效应）[5]。这种双向评级体系较 AAP 原有体系更为精细，后者很大程度上依赖研究设计层级。虽然 RCT 通常比观察性研究的评级更高，但这种对非 RCT 研究的考量能力在儿科领域尤为有益，因为儿科 RCT 相对少见。此外，GRADE 并非针对单项研究评价证据确定性，而是针对每一个结局进行评价。因此，同一项研究中的不同结局或多项研究中的同一个结局可能获得不同的质量评级。

指南专家组通过“证据至决策”框架制定临床推荐意见，该框架在权衡某一诊断方法或干预措施的理想效应与不理想效应的同时，还纳入了其他关键决策考量因素 [7]。在确定推荐意见的方向和强度时，GRADE 综合考虑获益与危害的平衡、证据确定性、患者价值观和偏好，以及成本、可接受性、可行性和公平性等次要因素。事实上，GRADE 框架明确允许在特定群体可能处于不利地位或干预措施效果存在差异的情况下，纳入与健康公平性相关的证据 [8]。所有这些考量因素决定了推荐意见的最终方向（“支持”或“反对”）和强度（强或条件性）[5,9]。这种二级体系旨在为最终使用者简化推荐意见。在新的《体重增长不佳 CPG》中，专家组针对无局灶性体征的患儿，基于获益甚微、中等程度危害、成本和降低公平性的考量，提出了一项反对早期进行常规诊断性检查的强推荐 [10]。

GRADE 是一个实用系统，用于处理关于替代干预措施或管理方法的临床问题，并非旨在消除判断或分歧，而是使这些分歧具体化和透明化 [5]。

GRADE 的潜在局限性

GRADE 系统存在局限性。采用 GRADE 制定的儿科 CPG 很可能具有较低的确定性评级和较少的强推荐，因此可能造成一种错误印象，即推荐意见缺乏证据支持，尽管确实是基于合理的结论和专家共识。GRADE 系统会降级间接证据，并且在将亚组变异性纳入证据评价时，有时会面临困难。在儿科领域，由于研究经费分配以及儿童作为受试者的特殊伦理情况，可能并不总是有大量基于 RCT 的证据可供制定指南。然而，GRADE 为评价 RCT 之外的各种证据类型和评估确定性提供了明确标准，同时兼顾了价值观和偏好、资源利用及公平性等背景因素。这种方法在儿科中尤具优势。

尽管 GRADE 系统采用高度结构化的方法，但 2 个不同的专家组基于相同的证据基础也可能得出不同的结论。例如，如果某项治疗的各研究效应量不一致，但所有点估计值均提示获益，则评审者可能在是否因不一致性而导致降级方面产生分歧，特别是当不确定性仅涉及治疗效应的量级时。评级差

异可能在过程中累积，并最终需要综合考量，这反映了在证据确定性本质为连续谱的情况下使用离散类别的挑战。在临界案例中，解释决策理由尤为重要[11]。如前所述，GRADE 并不消除分歧，而是使差异透明化。尽管如此，近期一项定性研究对使用 GRADE 做出决策时所存在的感知主观性提出了担忧[12]。即便使用 GRADE 这样的形式化系统，完全从证据解释中剔除个体判断也几乎是不可能的。

最后，GRADE 系统全面且严谨，但并不简单，需要充分的培训才能正确使用[12,13]。在过去 20 年中，GRADE 系统变得日益复杂，促使相关方努力简化其应用[14]。GRADE 的复杂性带来了更高的透明度，但也可能使其实施更加困难[15]。

结　论

AAP 向 GRADE 的过渡代表了 CPG 证据评价标准化进程的又一步骤，同时纳入了价值观和偏好、资源利用及公平性等背景因素，这体现了 AAP 致力于采用公认的工具以促进最佳实践，同时认识到患者和照护者需要理解这一过程，并基于高质量证据做出决策。AAP 必须监测这些变化对最终使用者 —— 即依赖 AAP 的 CPG 指导日常工作的儿科医师及其他临床医师 —— 产生的实际影响。当推荐意见在临床上有用、及时且基于现有最佳证据时，临床医师、患者和家庭都将从中获益。

缩略语

AAP：美国儿科学会

CPG：临床实践指南

GRADE：推荐分级的评估、制定与评价

RCT：随机对照试验

（蒋吴君 译　郝创利 校）

联系人：丁欣 etyxcbzx@suda.edu.cn

参考文献

1. Kredo T, Bernhardsson S, Machingaidze S, et al. Guide to clinical practice guidelines: the current state of play. Int J Qual Health Care. 2016;28(1):122–128. PubMed doi: 10.1093/intqhc/mzv115
2. Baker A, Young K, Potter J, Madan I. A review of grading systems for evidence-based guidelines produced by medical specialties. Clin Med (Lond). 2010;10(4):358–363. PubMed doi: 10.7861/clinmedicine.10-4-358
3. American Academy of Pediatrics Steering Committee on Quality Improvement and Management. Classifying recommendations for clinical practice guidelines. Pediatrics. 2004;114(3):874–877. PubMed doi: 10.1542/peds.2004-1260
4. Isaac A, Saginur M, Hartling L, Robinson JL. Quality of reporting and evidence in American Academy of Pediatrics guidelines. Pediatrics. 2013;131(4):732–738. PubMed doi: 10.1542/peds.2012-2027
5. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64(4):383–394. PubMed doi: 10.1016/j.jclinepi.2010.04.026
6. Atkins D, Eccles M, Flottorp S, et al; GRADE Working Group. Systems for grading the quality of evidence and the strength of recommendations I: critical appraisal of existing approaches The GRADE Working Group. BMC Health Serv Res. 2004;4(1):38. PubMed doi: 10.1186/1472-6963-4-38
7. Alonso-Coello P, Schünemann HJ, Moberg J, et al; GRADE Working Group. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ. 2016;353:i2016. PubMed doi: 10.1136/bmj.i2016
8. Welch VA, Akl EA, Guyatt G, et al. GRADE equity guidelines 1: considering health equity in GRADE guideline development: introduction and rationale. J Clin Epidemiol. 2017;90:59–67. PubMed doi: 10.1016/j.jclinepi.2017.01.014
9. Andrews JC, Schünemann HJ, Oxman AD, et al. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. J Clin Epidemiol. 2013;66(7):726–735. PubMed doi: 10.1016/j.jclinepi.2013.02.003
10. Kersten HB, Goday PS, Abdelhadi R, et al; AMERICAN ACADEMY OF PEDIATRICS and NORTH AMERICAN SOCIETY FOR PEDIATRIC GASTROENTEROLOGY, HEPATOLOGY AND NUTRITION. Clinical practice guideline for diagnosis and management of faltering weight. Pediatrics. 2026:e2025075764. PubMed doi: 10.1542/peds.2025-075764
11. Guyatt G, Oxman AD, Sultan S, et al. GRADE guidelines: 11. Making an overall rating of confidence in effect estimates

for a single outcome and for all outcomes. J Clin Epidemiol. 2013;66(2):151–157. PubMed doi: 10.1016/j.jclinepi.2012.01.006

12. Andrić A, Poklepović Perić T, Tokalić R, Marušić A, Puljak L. Experiences and challenges faced by systematic review authors in using the GRADE approach: a qualitative study. J Clin Epidemiol. 2025;185:111870. PubMed doi: 10.1016/j.jclinepi.2025.111870
13. Yan S, Guyatt GH, Ren X, et al. Evaluation of GRADE approach adherence in clinical practice guidelines: a systematic survey of criteria and their applications. J Clin Epidemiol. Published online February 2026:112186. doi: 10.1016/j.jclinepi.2026.112186
14. Colunga-Lozano LE, Wang Y, Agoritsas T, et al. Core GRADE unpacked: a summary of recent innovations in complementary GRADE methodology. J Clin Epidemiol. 2026;189:112047. PubMed doi: 10.1016/j.jclinepi.2025.112047
15. Atkins D, Best D, Briss PA, et al; GRADE Working Group. Grading quality of evidence and strength of recommendations. BMJ. 2004;328(7454):1490–1494. PubMed doi: 10.1136/bmj.328.7454.1490

【英文原件请参阅 *PEDIATRICS 2026;157(6):e2026076208*】