

Compreendendo a avaliação da qualidade dos estudos pela Classificação de Recomendações, Avaliação, Desenvolvimento e Análises (GRADE): da evidência às recomendações

Understanding study quality assessment through the Grading of Recommendations, Assessment, Development and Evaluations (GRADE): from evidence to recommendations

Leonardo Roeber¹, Mansueto Gomes Neto², André Rodrigues Durães², Paulo Eduardo Ocke Reis³, Priscila Pollo-Flores³, Rose Mary Lisboa da Silva⁴, Elmiro Santos Resende¹

RESUMO

A saúde baseada em evidências refere-se ao cuidadoso e preciso uso do conhecimento científico existente oriundo de pesquisas clínicas, reduzindo assim as incertezas no julgamento clínico frente ao paciente. As diretrizes destinadas a fornecer recomendações claras, devem seguir um conjunto de metodologias específicas. O objetivo deste artigo é descrever a metodologia e a aplicação da Classificação de Recomendações, Avaliação, Desenvolvimento e Análises como ferramenta fundamental neste delicado processo científico.

Descritores: Abordagem GRADE; Prática clínica baseada em evidências; Epidemiologia

ABSTRACT

Evidence-based healthcare is the careful and accurate use of scientific knowledge arising from clinical research, which reduces uncertainties regarding the medical judgement for the patient. As these guidelines intend to provide clear recommendations, they shall follow a set of specific methodologies. This study aims at describing the methodology and the application of the Grading of Recommendations, Assessment, Development and Evaluations as a critical tool in this intricate scientific process.

Keywords: GRADE approach; Evidence-based practice; Epidemiology

INTRODUÇÃO

A saúde baseada em evidências (SBE) refere-se ao uso criterioso, cuidadoso e preciso do conhecimento científico existente oriundo de pesquisas clínicas, utilizando-se de metodologias específicas, que garantam solidez e clareza nas informações a serem aplicadas na tomada de decisão clínica, reduzindo, assim, as incertezas na tomada de decisão junto ao paciente.⁽¹⁻³⁾

A Classificação de Recomendações, Avaliação, Desenvolvimento e Análises (GRADE, do inglês *Grading of Recommendations, Assessment, Development and Evaluations*) fornece critérios de qualidade adicionais, o que pode reduzir ou aumentar a qualidade das evidências para cada estudo avaliado.⁽⁴⁻¹²⁾ Limitações metodológicas, como presença de uma proporção de pacientes perdidos no acompanhamento, falta de cegamento em

¹Universidade Federal de Uberlândia, Uberlândia, MG, Brasil.

²Universidade Federal da Bahia, Salvador, BA, Brasil.

³Universidade Federal Fluminense, Niterói, RJ, Brasil.

⁴Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil.

Data de submissão: 31/07/2019. **Data de aceite:** 18/08/2019.

Autor correspondente: Leonardo Roeber. Avenida Pará, 1.720 – Umarama – CEP: 38400-902 – Uberlândia, MG, Brasil.

Tel.: (34) 98803-9878 – E-mail: leonardoroerber@hotmail.com

Fonte de auxílio à pesquisa: nenhum. **Conflitos de interesse:** nenhum.

Contribuição dos autores: Concepção e delineamento do projeto: LR, MGN e ARD.

Coleta, análise e interpretação de dados: LR, PEOR, PPF e RMLS.

Redação e revisão crítica do manuscrito: LR, MGN, ARD, PEOR, PPF, RMLS e ESR.

Aprovação da versão final do manuscrito a ser publicado: LR, MGN, ARD, PEOR, PPF, RMLS e ESR.

um estudo randomizado ou vários outros vieses, são responsáveis pela redução da qualidade da evidência.^(1,13,14) Em situações nas quais diferentes critérios levariam a diferentes recomendações, a classificação de evidência clínica deve estar relacionada com o limiar da questão clínica (Quadro 1).⁽¹³⁾

DA EVIDÊNCIA ÀS RECOMENDAÇÕES

Cada recomendação baseia-se na evidência derivada dos resultados de questões clínicas. Quando a qualidade da evidência difere entre os resultados, a GRADE estabelece que o grau de qualidade das evidências determine a qualidade geral delas.^(1,2)

Consequentemente, a qualidade geral da evidência seria moderada.⁽³⁾ Ao usar as recomendações da GRADE, os profissionais de saúde têm necessidade de compreender a força da recomendação e da classificação da certeza da evidência. Além disso, informações adicionais cruciais, como observações, efeitos completos das alternativas de tratamento, ou considerações sobre valores e preferências e custo, devem estar prontamente acessível. Recomendações desenvolvidas com a abordagem da GRADE são classificadas como forte ou fraca. Uma forte recomendação reflete a confiança do painel de especialistas que uma consequência da ação proposta supera ou não os indesejáveis efeitos.⁽¹⁵⁾

A superação dos efeitos desejáveis de uma intervenção sobre efeitos indesejáveis reflete a força de uma recomendação. Recomendações fortes significam que a maioria dos pacientes informados escolheria o tratamento recomendado, e os profissionais de saúde podem estruturar suas interações com os pacientes em conformidade.⁽¹⁶⁾ O fraseado usual para recomendações fortes é “recomendamos”. Uma recomendação “fraca” reflete um equilíbrio ténue entre benefícios e desvantagens, incerteza quanto à magnitude das vantagens e desvantagens, mudança ou alta variabilidade nos valores e preferências de pacientes, além de ausência adequada de uma justificativa de custo da intervenção planejada.⁽¹⁵⁾

Recomendações fracas significam que as escolhas dos pacientes variam de acordo com seus valores e pre-

ferências, e os profissionais de saúde devem assegurar que o atendimento ao paciente esteja de acordo com valores e preferências.⁽¹⁶⁾ Uma recomendação fraca é geralmente como “sugerimos” ou “pode ser considerado” (Quadro 2). Tais informações podem ser encontradas no resumo das tabelas, que fornecem a taxa para todos os resultados relevantes e os valores numéricos para as estimativas relativas e absolutas de o efeito.⁽¹⁵⁾

A força de uma recomendação é determinada pelo equilíbrio entre consequências de estratégias alternativas de manejo desejáveis e indesejáveis, qualidade de evidências, variabilidade de valores e preferências e uso de recursos.⁽¹⁶⁾ Recomenda-se que os especialistas considerem o tipo de evidências que eles avaliarão e também a experiência de pesquisa dos avaliadores, antes de selecionar uma avaliação crítica e um sistema de classificação. Se a literatura a ser revisada engloba uma mistura mais ampla de metodologias, incluindo análises qualitativas e sociológicas ou evidências baseadas em opinião, bem como pesquisa quantitativa, é recomendado que todos os avaliadores tivessem treinamento em avaliação e classificar provas usando o sistema a ser empregado para a diretriz (Quadro 4).⁽¹⁸⁾

Pelo contrário, a GRADE considera aspectos adicionais, ao fazer recomendações, mas eles não fazem parte da qualidade do julgamento das evidências. Por outro lado, a classificação se concentra menos no tipo de estudo e considera múltiplos aspectos dos dados disponíveis, ao avaliar a qualidade. No entanto, a GRADE é demorada, porque os usuários devem considerar múltiplos aspectos durante a avaliação de qualidade, cada um dos quais abertos a algum grau de subjetividade. Portanto, uma crítica que poderia ser relatada, legitimamente, para a GRADE é que parece que foi desenvolvida por acadêmicos para acadêmicos. Se a maior complexidade da GRADE, em comparação com outros métodos, resulta no fracasso dos desenvolvedores de diretrizes clínicas em entender o processo, então isso limitará a qualidade de alcançar o objetivo para ser padronizado. Uma solução potencial é produzir métodos alternativos de exibição de dados para diretrizes mais fáceis de usar voltados aos profissionais de saúde, que são os usuários finais.

Quadro 1. A avaliação da qualidade da evidência, segundo a Classificação de Recomendações, Avaliação, Desenvolvimento e Análises.

Qualidade inicial da evidência	Desenho de estudo	Abaixar a qualidade se	Aumentar a qualidade se
Alta	Ensaio clínico randomizado, revisão sistemática, metanálise	Limitações do estudo Inconsistência	Magnitude do efeito Gradiente dose-resposta
Moderada	Ensaio clínico randomizado, revisão sistemática, metanálise com problemas na condução	Indireção Impressão Viés de publicação	Todos os confundidores plausíveis Ter efeito reduzido
Baixa	Estudo observacional (estudo de coorte, estudo de caso-controle)		
Muito baixa	Qualquer outra evidência (série de casos)		

Quadro 2. Critérios para classificar os Níveis de Evidência e força das recomendações, de acordo com os critérios da Classificação de Recomendações, Avaliação, Desenvolvimento e Análises.

Grau de recomendação	Forte recomendação	Recomendação moderada	Recomendação fraca	Recomendação negativa
Conclusão de evidência	Benefícios mais fortes que risco	Benefícios maiores que o risco	Benefícios iguais que risco	Sem benefícios
Alto Nível de Evidência	Recomendação forte baseada em alto Nível de Evidência	Recomendação moderada com base no alto nível de Evidência	Recomendação fraca com base no alto Nível de Evidência	Nenhuma recomendação com base no alto Nível de Evidência
Nível moderado/baixo de evidência	Recomendação forte baseada em nível moderado/baixo de evidência	Recomendação moderada baseada em moderado/baixo Nível de Evidência	Recomendação fraca baseada em moderado/baixo Nível de Evidência	Nenhuma recomendação baseada em moderado/baixo Nível de Evidência
Nível muito baixo de evidência	Recomendação forte baseada na opinião de especialistas	Recomendação moderada com base no baixíssimo Nível de Evidência	Recomendação fraca com base no baixíssimo Nível de Evidência	Nenhuma recomendação com base no baixíssimo Nível de Evidência
Relato das recomendações	“Nós recomendamos” “Deveríamos” “É recomendado/ indicado/útil/eficaz”	“Nós sugerimos” “É provavelmente recomendado” “Pode ser útil ou eficaz”	“Nós podemos sugerir” “Pode ser considerado”	“Nós não recomendamos” “Não é útil/eficaz”

GRADE é um sistema abrangente, objetivo e robusto, que exige reprodutibilidade suficiente para ser usado todos os dias na prática, mesmo por painéis de especialistas com conhecimento relativamente limitado à metodologia de pesquisa.⁽¹⁷⁾ No entanto, essa classificação tem sido desenvolvida para abordar questões sobre gerenciamento alternativo, estratégias, intervenções ou políticas. Não foi preparada para perguntas sobre risco ou prognóstico. Preparar uma diretriz implica várias etapas da GRADE. Por esmagador consenso, a GRADE provou seu mérito na avaliação do tratamento e de intervenções, além da abordagem clínica (Quadro 3).

Quadro 3. Fatores que afetam a força da recomendação

Fator	Explicação
Equilíbrio entre desejável e efeitos indesejáveis	Quanto mais significativa a diferença entre os efeitos desejáveis e indesejáveis, maior a garantia de verossimilhança/recomendação firme Quanto mais estreito o gradiente, maior a probabilidade de uma recomendação condicional/fraca
Qualidade de evidência	Quanto maior a qualidade da evidência, maior a probabilidade de uma recomendação substancial
Valores e preferências	Quanto mais valores e preferências variam, ou quanto maior a incerteza em valores e preferências, maior a verossimilhança/recomendação condicional/fraca
Custos (alocação de recursos)	Quanto maiores os custos de intervenção (maiores os recursos consumidos), menor a probabilidade de uma substancial recomendação

Quadro 4. Implicações de recomendações fortes e condicionais/fracas para diferentes usuários de diretrizes.

Usuário	Forte	Condicional ou fraco
Pacientes	A maioria desejaria a ação recomendada Apenas uma pequena proporção não	A maioria gostaria que a ação fosse sugerida Muitos não
Profissionais de saúde	A adesão a essa recomendação pode ser usada como critério de qualidade ou indicador de desempenho Ajuda formal à decisão não é necessária para auxiliar indivíduos a tomar decisões consistentes com valores e preferências	As ajudas de decisão ajudam os indivíduos a tomar decisões consistentes com seus valores e preferências
Formuladores de políticas	Recomendação adotada como política na maioria das situações, incluindo para o uso como indicadores de desempenho	Debates substanciais e envolvimento de muitas partes interessadas Os indicadores de desempenho se concentram em deliberações adequadas sobre opções de gerenciamento

Os profissionais de saúde podem achar as diretrizes menos confusas e mais simples de assimilar se apenas o resultado da avaliação da qualidade for apresentado, tendo entendido que essa recomendação foi definida após uma avaliação rigorosa, usando a abordagem GRADE.⁽¹⁸⁾

CONCLUSÃO

A GRADE é um critério transparente, explícito e abrangente para graduação da qualidade da evidência e separa essa classificação da e as recomendações. No entanto, é um método complexo para não investigadores e profissionais de saúde de forma geral. Além disso, existe um grau de subjetividade que pode interferir nas recomendações e diretrizes. Existe um fator não controlado, relativo à não publicação de estudos, com resultados negativos ou ausência de resultados que não deve ser negligenciado, muito embora não seja passível de resolução. Apesar das ressalvas, a GRADE é a melhor ferramenta que temos atualmente para avaliar a tomada de condutas baseada em estudos científicos disponíveis na literatura.

REFERÊNCIAS

- Bertolaccini L, Rocco G. The forest of methodology and the writing of evidence-based medicine papers. *Eur J Cardiothorac Surg*. 2018;54(4):615-21. doi: 10.1093/ejcts/ezy264
- Brasil. Ministério da Saúde. Secretaria de Ciência, Tecnologia e Insumos Estratégicos. Departamento de Ciência e Tecnologia. Coordenação-Geral de Gestão do Conhecimento. Diretrizes metodológicas: Sistema GRADE – Manual de graduação da qualidade da evidência e força de recomendação para tomada de decisão em saúde [Internet]. Brasília, DF: Ministério da Saúde; 2014 [citado 2020 Nov 26]. Disponível em: <https://pesquisa.bvsalud.org/bvsms/resource/pt/mis-37111>
- Brozek JL, Akl EA, Alonso-Coello P, Lang D, Jaeschke R, Williams JW, Phillips B, Lelgemann M, Lethaby A, Bousquet J, Guyatt GH, Schünemann HJ; GRADE Working Group. Grading quality of evidence and strength of recommendations in clinical practice guidelines. Part 1 of 3. An overview of the GRADE approach and grading quality of evidence about interventions. *Allergy*. 2009;64(5):669-77. doi: 10.1111/j.1398-9995.2009.01973.x
- Brozek JL, Akl EA, Compalati E, Kreis J, Terracciano L, Focchi A, Ueffing E, Andrews J, Alonso-Coello P, Meerpohl JJ, Lang DM, Jaeschke R, Williams JW Jr, Phillips B, Lethaby A, Bossuyt P, Glasziou P, Helfand M, Watine J, Afilalo M, Welch V, Montedori A, Abraha I, Horvath AR, Bousquet J, Guyatt GH, Schünemann HJ; GRADE Working Group. Grading quality of evidence and strength of recommendations in clinical practice guidelines part 3 of 3. The GRADE approach to developing recommendations. *Allergy*. 2011;66(5):588-95. doi: 10.1111/j.1398-9995.2010.02530.x
- Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y et al. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *J Clin Epidemiol*. 2013;66(7):719-25. doi: 10.1016/j.jclinepi.2012.03.013
- Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol*. 2011;64(4):395-400. doi: 10.1016/j.jclinepi.2010.09.012
- Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). *J Clin Epidemiol*. 2011;64(4):407-15. doi: 10.1016/j.jclinepi.2010.07.017
- Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, et al. GRADE guidelines: 5. Rating the quality of evidence—publication bias. *J Clin Epidemiol*. 2011;64(12):1277-82. doi: 10.1016/j.jclinepi.2011.01.011
- Puhan MA, Schünemann HJ, Murad MH, Li T, Brignardello-Petersen R, Singh JA, Kessels AG, Guyatt GH; GRADE Working Group. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. *BMJ*. 2014;349:g5630. doi: 10.1136/bmj.g5630. Erratum in: *BMJ*. 2015;350:h3326
- Iorio A, Spencer FA, Falavigna M, Alba C, Lang E, Burnand B, McGinn T, Hayden J, Williams K, Shea B, Wolff R, Kujpers T, Perel P, Vandvik PO, Glasziou P, Schünemann H, Guyatt G. Use of GRADE for assessment of evidence about prognosis: rating confidence in estimates of event rates in broad categories of patients. *BMJ*. 2015;350:h870. doi: 10.1136/bmj.h870
- Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol*. 2011;64(4):401-6. doi: 10.1016/j.jclinepi.2010.07.015
- Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, DeBeer H, Jaeschke R, Rind D, Meerpohl J, Dahm P, Schünemann HJ. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011;64(4):383-94. doi: 10.1016/j.jclinepi.2010.04.026
- Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, Devereaux PJ, Montori VM, Freyschuss B, Vist G, Jaeschke R, Williams JW Jr, Murad MH, Sinclair D, Falck-Ytter Y, Meerpohl J, Whittington C, Thorlund K, Andrews J, Schünemann HJ. GRADE guidelines 6. Rating the quality of evidence—imprecision. *J Clin Epidemiol*. 2011;64(12):1283-93. doi: 10.1016/j.jclinepi.2011.01.012
- Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Glasziou P, Jaeschke R, Akl EA, Norris S, Vist G, Dahm P, Shukla VK, Higgins J, Falck-Ytter Y, Schünemann HJ; GRADE Working Group. GRADE guidelines: 7. Rating the quality of evidence—inconsistency. *J Clin Epidemiol*. 2011;64(12):1294-302. doi: 10.1016/j.jclinepi.2011.03.017
- Neumann I, Santesso N, Akl EA, Rind DM, Vandvik PO, Alonso-Coello P, Agoritsas T, Mustafa RA, Alexander PE, Schünemann H, Guyatt GH. A guide for health professionals to interpret and use recommendations in guidelines developed with the GRADE approach. *J Clin Epidemiol*. 2016;72:45-55. doi: 10.1016/j.jclinepi.2015.11.017
- Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, Schünemann HJ; GRADE Working Group. Going from evidence to recommendations. *BMJ*. 2008;336(7652):1049-51. doi: 10.1136/bmj.39493.646875.AE. Erratum in: *BMJ*. 2008;336(7658):doi:10.1136/bmj.a402
- Kumar A, Miladinovic B, Guyatt GH, Schünemann HJ, Djulbegovic B. GRADE guidelines system is reproducible when instructions are clearly operationalized even among the guidelines panel members with limited experience with GRADE. *J Clin Epidemiol*. 2016;75:115-8. doi: 10.1016/j.jclinepi.2015.11.020
- Baker A, Potter J, Young K, Madan I. The applicability of grading systems for guidelines. *J Eval Clin Pract*. 2011;17(4):758-62. doi: 10.1111/j.1365-2753.2011.01693.x
- Cooper NA, Khan KS, Clark TJ. Evidence quality in clinical guidelines: a comparison of two methods. *Acta Obstet Gynecol Scand*. 2015;94(12):1283-9. doi: 10.1111/aogs.12759
- Schünemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, Williams JW Jr, Kunz R, Craig J, Montori VM, Bossuyt P, Guyatt GH; GRADE Working Group. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ*. 2008;336(7653):1106-10. doi: 10.1136/bmj.39500.677199. AE. Erratum in: *BMJ*. 2008;336(7654). doi: 10.1136/bmj.a139. Schünemann, A Holger J [corrected to Schünemann, Holger J]