

Fisioter Bras 2017;18(5):657-66

REVISÃO

GRADE: um sistema para graduar qualidade de evidência e força da recomendação e as implicações para a prática fisioterapêutica

GRADE: a system to graduate quality of evidence and strength of recommendation and the implications to the physical therapy practice

Andrea Lemos, Ft., D.Sc.

Docente do Departamento de Fisioterapia da Universidade Federal de Pernambuco (UFPE)

Recebido em 15 de maio de 2017; aceito em 8 de outubro de 2017.

Endereço de correspondência: Andrea Lemos, Departamento de Fisioterapia, Av. Professor Moraes Rego, 1235, Cidade Universitária, 50670-901 Recife PE, E-mail: andrealemos4@gmail.com

Resumo

Introdução: O sistema de graduação de qualidade de evidência e força de recomendação, denominado *Grading of Recommendations, Assessment, Development and Evaluation* (GRADE), tem sido adotado por organizações para o desenvolvimento de diretrizes e por revisores sistemáticos para análise da qualidade da evidência. Esse artigo tem como objetivo descrever os princípios e a abordagem do sistema GRADE. **Métodos:** Revisão de literatura utilizando as bases de dados Medline, CINAHL e Web of Science. **Resultados:** O sistema classifica a qualidade da evidência em alta, moderada, baixa e muito baixa, de acordo com fatores que consideram: a metodologia dos estudos, a consistência e a precisão dos resultados, o direcionamento da evidência e o viés de publicação. A força da recomendação é graduada em forte ou fraca e baseia-se não apenas na qualidade da evidência, mas no equilíbrio entre benefício e malefício da intervenção, valores e preferências do paciente e utilização dos recursos e custos. Os achados apresentados por esse sistema refletem a extensão da confiabilidade de que as estimativas dos efeitos encontradas estão corretas. **Conclusão:** O entendimento do sistema GRADE facilita a interpretação crítica de revisões sistemáticas e diretrizes que o utilizam e espera-se que haja um aumento da sua incorporação na produção desses estudos na área da fisioterapia.

Palavras-chave: prática clínica baseada em evidências, revisão, guia de prática clínica.

Abstract

Introduction: The grading system of quality of evidence and strength of recommendation, called *Grading of Recommendations, Assessment, Development and Evaluation* (GRADE) has been adopted by organizations to develop guidelines and by systematic reviewers to analyze the quality of evidence. **Methods:** Literature review using Medline, CINAHL e Web of Science databases. This paper aimed to describe the principles and objectives of the GRADE system. **Results:** The system classifies quality of evidence as high, moderate, low and very low, according to factors that consider: the study's methodology, consistency and precision of the results, directness of the evidence and publication bias. The strength of recommendation is graduated in strong or weak and it is based not only on the quality of evidence, but on the balance between benefit and harm of the intervention, values and preferences of the patients and resources and costs use. The findings presented by this system reflect the extent of the confidence that the effect estimates are correct. **Conclusion:** Understanding the GRADE system facilitates the critical interpretation of systematic reviews and guidelines that use it and it is expected that its incorporation in the production of these studies in the area of physical therapy increases.

Key-words: evidence-based practice, review, practice guideline.

Introdução

O sistema GRADE (*Grades of Recommendation, Assessment, Development and Evaluation*) surgiu em 2000, através de uma colaboração de pesquisadores, epidemiologistas e estatísticos que tinham como objetivo uniformizar os sistemas de graduação dos níveis de

evidência e recomendações existentes. Esse sistema apresenta como proposta graduar a qualidade de evidência em revisões sistemáticas e diretrizes e graduar a força das recomendações nas diretrizes [1].

O primeiro artigo publicado pelo grupo, intitulado: *Grading quality of evidence and strength of recommendations*, surge em 2004, no *British Medical Journal* (BMJ). Nesta publicação é apresentada a proposta inicial do GRADE quanto às definições de qualidade de evidência, força da recomendação e o diferencial em relação aos outros sistemas de avaliação de níveis de evidência vigentes [2]. Em 2006 o *British Medical Journal* (BMJ) dispõe como requisito, nas instruções aos autores, o uso do sistema GRADE para graduar a evidência dos *Guidelines* submetidos à revista [3]. Dois anos mais tarde, em 2008, o grupo lança uma série de seis artigos apresentando a proposta do GRADE [4-9] e entre 2011 e 2013 é publicada outra série de 15 artigos, pelo *Journal of Clinical Epidemiology*, esmiuçando toda a abordagem proposta pelo sistema em detalhes [10,11].

Atualmente o sistema GRADE vem sendo utilizado por mais de 60 organizações e Sociedades internacionais, entre elas: a Organização Mundial de Saúde, colaboração Cochrane, Sociedade Torácica Americana, Sociedade Endócrina Americana, Sociedade Respiratória Européia, Colégio Americano de Médicos Torácicos, Sociedade de doença infecciosa da América e *Agency for Health Care Research and Quality* (AHRQ) dos Estados Unidos. Importantes Institutos de Diretrizes internacionais também incorporaram esse sistema como, por exemplo, o *National Institutes of Health and Clinical Excellence* (NICE) e o *Scottish Intercollegiate Guideline Network* (SIGN) e ele vem sendo instituído, de maneira crescente, em diversas revisões sistemáticas recentemente publicadas [12-15]. Mais de 200 membros compõe o quadro atual e estão envolvidos no constante aprimoramento dessa abordagem [10]. Além das publicações existentes, é possível também adquirir informações através de um site organizado pelo grupo de trabalho, além de um *Handbook* disponível na internet [12,14].

No cenário da Fisioterapia, algumas iniciativas foram dadas, tanto para introduzir os conceitos desta abordagem como para encorajar o seu uso no desenvolvimento de diretrizes e revisões sistemáticas, em revistas importantes na área como a *Physical Therapy* e o *Journal of Physiotherapy* [15,16]. Diante deste contexto, o presente artigo tem como objetivo descrever os princípios do sistema GRADE e sua abordagem para graduar a qualidade de evidência e formular as recomendações para a prática clínica, com o intuito de atualizar os fisioterapeutas na leitura crítica da produção científica usuária do GRADE.

Métodologia

Foi realizada uma revisão da literatura nas bases de dados *Medline* (via PUBMED), CINHAL (via EBSCO) e *Web of Science*, no período entre 2000 (ano em que o sistema foi criado) e 2016. Além da consulta no site oficial do sistema GRADE (<http://www.gradeworkinggroup.org/>) [12] e no seu *Handbook* [14]. Procurou-se por artigos metodológicos relacionados ao desenvolvimento e operacionalização do sistema GRADE e não houve restrição linguística.

Resultados e discussão

Definições

O GRADE corresponde a um sistema de graduação para qualidade de evidência em revisões sistemáticas e diretrizes e recomendações nas diretrizes. Esse sistema oferece um processo estruturado e transparente na apresentação sumária da evidência [17].

O sistema GRADE apresenta duas definições para “qualidade de evidência”. Uma para a revisão sistemática e outra para o contexto das diretrizes. Para revisão sistemática, “qualidade de evidência” reflete a extensão da nossa confiabilidade de que as estimativas do efeito estão corretas. Para as diretrizes reflete a extensão da nossa confiabilidade que as estimativas do efeito estão adequadas para suportar uma decisão particular ou recomendação [17]. Esse sistema defende a confiança na estimativa do efeito. Portanto, a utilidade da estimativa da magnitude da intervenção ou efeito depende na confiabilidade desta estimativa [4].

O GRADE apresenta uma mudança no raciocínio, uma vez que separa qualidade de evidência e graus de recomendação, ou seja, alto nível de evidência não necessariamente

implica forte recomendação e uma forte recomendação pode surgir de um nível de evidência baixo [4]. Desta forma, estudos observacionais rigorosos podem proporcionar uma evidência mais forte do que uma série de casos não controlada.

Desde que o GRADE foi proposto, o grupo de trabalho vem demonstrando as vantagens em relação aos outros sistemas como a separação clara entre a qualidade de evidência e a força da recomendação; a transparência nos critérios de graduação na qualidade da evidência; a ênfase na escolha dos desfechos; o reconhecimento dos valores e preferências do paciente nas recomendações; a apresentação dos resultados da evidência através de tabelas e a facilidade na interpretação dos termos utilizados na recomendação [2,10].

Classificação da qualidade da evidência

O GRADE classifica a qualidade da evidência em quatro níveis: alto, moderado, baixo e muito baixo (Figura 1). A evidência baseada em ensaios clínicos randomizados e controlados inicia como alta qualidade, podendo diminuir, enquanto que os estudos observacionais iniciam com uma classificação baixa, podendo aumentar o nível da qualidade [17].

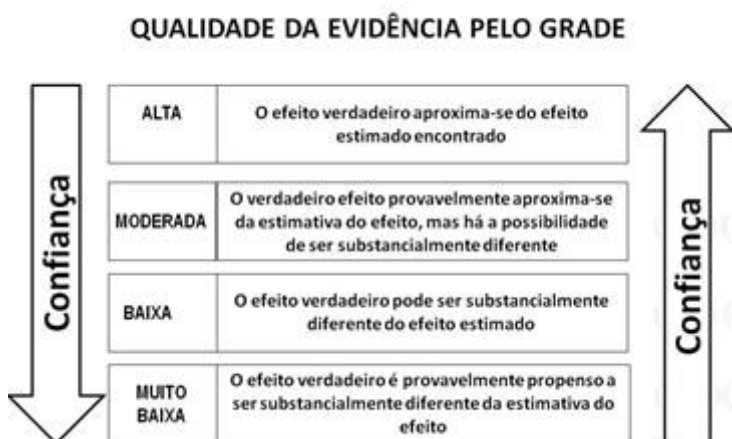


Figura 1 - Classificação da qualidade da evidência de acordo com o Sistema Grading of Recommendations, Assessment, Development and Evaluation (GRADE) e a confiança na estimativa do efeito encontrado.

Cinco são os fatores que podem diminuir a evidência nos ensaios clínicos: limitação do estudo (study limitation), inconsistência (inconsistency), imprecisão (imprecision), direcionamento (indirectness) e viés de publicação (publication bias). É importante destacar a semântica dos termos utilizados. O uso do prefixo latino “in, im”, indicando negação/sentido contrário, mostra que está sendo mensurado o grau de imperfeição dos estudos. Confia-se, portanto, no estudo menos inconsistente, menos impreciso, menos indireto, deixando de qualificar os estudos em boa e má qualidade, como preconizava as primeiras escalas utilizadas para avaliar a qualidade dos estudos em revisões sistemáticas [5].

- **Limitação do Estudo:** corresponde à validade interna do estudo quanto aos critérios metodológicos utilizados para controlar os vieses do estudo. Este item considera os seguintes aspectos: randomização, sigilo de alocação, mascaramento, controle das perdas, descrição seletiva do desfecho (demonstração de um efeito de tratamento aparente e falha em apresentar outros desfechos que mostram nenhuma evidência de efeito), parada precoce do ensaio e uso de instrumento válido para mensurar os desfechos. A falta de controle desses procedimentos metodológicos implica queda na qualidade da evidência [18].
- **Inconsistência:** reflete as diferentes estimativas do efeito do tratamento como resultado da heterogeneidade ou variabilidade dos estudos. Essa variabilidade pode surgir de diferenças na população, intervenção ou nos desfechos. Quando há uma heterogeneidade e não existe uma explicação plausível, a qualidade da evidência

diminui. A variabilidade nos resultados dos estudos individuais pode surgir de diferenças clínicas na população (maior efeito em um subgrupo de pacientes), intervenção (maior efeito com doses maiores de medicamentos), desfechos (diferenças na definição de resposta ao tratamento) e de limitações metodológicas (problemas na randomização, ausência de mascaramento) [19].

- *Imprecisão*: reflete quão precisa é a estimativa do efeito do tratamento através do intervalo de confiança obtido. Amplos intervalos de confiança, que incluem tanto importantes efeitos benéficos quanto nenhum efeito, resultam em incerteza no resultado. Portanto, a imprecisão ocorre quando o intervalo de confiança é amplo, as amostras são pequenas ou há poucos eventos [20].
- *Direcionamento*: refere à extensão na qual a população, intervenção e desfecho são similares aos de interesse. Comparações indiretas surgem quando, por exemplo, a evidência vem de estudos que comparam intervenção A com placebo e intervenção B com placebo, mas não há uma comparação direta de A versus B. O GRADE também considera o direcionamento no escopo da recomendação, ou seja, a evidência que suporta a recomendação é também indireta quando é obtida de estudos nos quais a população, intervenção ou desfecho de interesse foram diferentes do proposto pela recomendação [21].
- *Viés de Publicação*: Quando a evidência é obtida de poucos estudos, com amostras pequenas, ou financiados pela indústria farmacêutica a qualidade da evidência será reduzida [22].

Desta forma, uma determinada evidência pode ser proveniente de limitações em mais de um dos fatores descritos acima e, desta forma, quanto mais sérias forem essas limitações menor será a qualidade de evidência (Tabela I).

Em relação aos estudos observacionais, o GRADE considera-os como baixo nível, podendo aumentar se houver uma grande magnitude de efeito do tratamento ou evidência de uma gradiente dose-resposta ou ainda se todos os vieses plausíveis diminuiria ou aumentaria a magnitude de um efeito demonstrada se nenhum efeito fosse observado [23] (Tabela I).

- *Grande magnitude de efeito*: o sistema GRADE define um grande efeito se houver um Risco Relativo (RR) maior que 2,0 ou menor que 0,5, baseado em evidência consistente de pelo menos dois estudos, sem confundidores plausíveis e um efeito muito grande com um RR maior que 5,0 ou menor que 0,2, também baseado em uma evidência direta sem vieses na validade interna [5,23].
- *Gradiente dose-resposta*: a presença de um gradiente dose-resposta ocorre quando níveis elevados de um determinado fator prognóstico (maior duração, maior intensidade, maior dosagem) proporcionam um aumento no tamanho do efeito, em comparação a níveis mais baixos desse fator. A presença desta relação pode aumentar a confiança nos resultados [5,23].
- *Confundidores plausíveis*: ocorre quando estudos observacionais falham em demonstrar uma associação, mas todos os vieses plausíveis teriam aumentado ou diminuído um efeito da intervenção. Essa é uma definição que requer uma interpretação mais rebuscada, uma vez que não se trata apenas de haver um controle dos fatores confundidores no estudo. A pretensão deste raciocínio metodológico implica acreditar nos resultados de um estudo observacional que não observou uma determinada associação que se esperaria devido aos plausíveis efeitos confundidores. O exemplo clássico citado pela literatura do GRADE é sobre a metanálise que encontrou uma maior proporção de mortes em hospitais privados com fins lucrativos comparados a hospitais privados sem fins lucrativos. Esperar-se-ia, neste caso, que devido aos vieses encontrados, quanto às diferenças na gravidade das doenças e do efeito dos pacientes com seguro, a estimativa de associação fosse a favor dos hospitais privados com fins lucrativos. No entanto, essa revisão sistemática, envolvendo 38 milhões de pacientes, falhou em demonstrar essa associação. Deste modo, pode-se elevar a evidência desses estudos observacionais, fornecendo mais confiança nesses resultados [23-25].

Tabela I - Fatores que influenciam a qualidade de evidência nos estudos (experimental versus observacional) e a graduação da evidência de acordo com o Sistema Grading of Recommendations, Assessment, Development and Evaluation (GRADE).

qualidade de evidência	
Fatores que podem diminuir (Ensaio clínico)	Fatores que podem aumentar (Estudos observacionais)
Limitações do desenho do estudo	Grande Magnitude do efeito
Inconsistência dos resultados	Gradiente Dose-resposta
Imprecisão dos resultados	Quando os confundidores plausíveis poderiam
Direcionamento da evidência	reduzir ou aumentar o efeito demonstrado se
Viés de publicação	nenhum efeito fosse observado
graduação em cada fator	
Sem limitação (não diminui)	Sem efeito (não aumenta)
Limitação séria (-1 nível)	Efeito grande (+1 nível)
Limitação muito séria (- 2 níveis)	Efeito muito grande (+ 2 níveis)

Qualidade de evidência global

O Grade propõe uma avaliação da qualidade de evidência por desfecho. Desta forma, cada desfecho investigado é avaliado separadamente em relação aos fatores acima descritos, dependendo do tipo de estudo, podendo diminuir ou aumentar a qualidade da evidência em um ou dois níveis se houver um comprometimento sério ou muito sério [26].

Esse sistema não fornece uma classificação quantitativa de risco de viés. A fonte de viés pode causar diferentes repercussões dependendo do desfecho. Assim, a qualidade de evidência pode mudar entre os desfechos de um mesmo estudo [26]. Exemplo: suponha que um determinado ensaio clínico objetivou avaliar a eficácia de uma intervenção fisioterapêutica em pacientes com esclerose múltipla. Este estudo apresentou como desfechos: “marcadores inflamatórios”, através de uma análise sanguínea, e a “função muscular” (mensurada por uma escala) e não houve mascaramento. Para o desfecho: “marcadores inflamatórios” a ausência do mascaramento interferirá pouco, mas para a análise da “função muscular” haverá maior influência, pois se trata de um desfecho subjetivo e o fato de o avaliador ter conhecimento do grupo tratado poderá influenciar no processo avaliativo.

É importante ressaltar que o julgamento do GRADE não se refere à avaliação individual de cada estudo, através da validade interna (qualidade metodológica) como era proposto pelos outros instrumentos de avaliação da evidência. De acordo com o GRADE, o corpo da evidência transcende o risco de viés e, como está relacionada à confiança na estimativa do efeito observado, considera outros fatores: imprecisão, inconsistência, direcionamento e viés de publicação. Desta forma, para que esse raciocínio seja eficiente, o sistema propõe aos revisores sistemáticos e aos organizadores de diretrizes a escolha de desfechos potencialmente importantes para o paciente para serem apresentados como a evidência da intervenção. O sistema propõe categorizar os desfechos em críticos, importantes, mas não críticos e de limitada importância [26]. Por exemplo, desfechos como mortalidade, qualidade de vida e funcionalidade podem ser considerados críticos, enquanto que força muscular, importante, mas não crítico e dados eletromiográficos de limitada importância para uma recomendação.

A escolha de desfechos críticos e importantes centrados no paciente desempenha um papel essencial para que a abordagem do GRADE seja aplicada e interpretada. Como a análise da evidência deixa de focar o estudo de forma global e passa para os desfechos, a evidência global é fornecida pelo desfecho crítico e importante clinicamente para o paciente. Deste modo, se o desfecho crítico e importante for classificado como de alta qualidade e os demais não se enquadrarem nesse nível, a qualidade de evidência global será de alta qualidade. Caso ocorra o oposto, ou seja, se o desfecho crítico e importante for considerado de baixa qualidade e os demais de alta qualidade, a evidência global será de baixa qualidade [26].

O GRADE proporciona o sumário da evidência através de tabelas denominadas perfil de evidência (*Evidence profile*- EP) e sumário de achados (*summary of findings* - SoFs). Essas tabelas apresentam diferentes propósitos. As tabelas “sumário de achados” – SoFs- apresentam a mesma informação das tabelas “perfil de evidência”, com exceção dos detalhes da qualidade da evidência e com adição de uma coluna de comentários [28-30].

- *Tabelas perfil de evidência (EP)* - essas tabelas apresentam os resultados relevantes para cada desfecho através das linhas. São expressos nas colunas: números de estudos, número de pacientes, o desenho do estudo, as comparações realizadas, a estimativa do efeito observado em termos relativos e absolutos com os respectivos intervalos de confiança a 95%. Ademais são expostos os fatores que graduarão a evidência (limitação dos estudos, inconsistência, imprecisão, direcionamento e viés de publicação). Por fim, é apresentada a qualidade da evidência para cada desfecho (alta, moderada, baixa ou muito baixa). Esse tipo de tabela deve ser utilizado pelos desenvolvedores de diretrizes e revisores sistemáticos para melhor interpretação da avaliação da evidência [28]. (Tabela II). Como exemplo, observa-se seu uso na revisão “*Some physiotherapy treatments may relieve menstrual pain in women with primary dysmenorrhea: a systematic review*” [31].
- *Tabelas Sumário de achados (SoFs)* - São mais sucintas e apresentam 7 elementos recomendados: desfechos importantes, o risco estimado no grupo controle, o risco no grupo de intervenção, a magnitude de efeito relativa, número de participantes e de estudos para cada desfecho, a qualidade da evidência do efeito estimado e comentários. Essa tabela deve ser utilizada nas publicações de revisões sistemáticas e [28] (Tabela III). Pode ser encontrada, como exemplo, na revisão da Cochrane intitulada “*Self management for patients with chronic obstructive pulmonary disease*” [32].

[Tabela II - Exemplo da tabela de perfil de evidência \(Evidence profile-EP\) do Sistema Grading of Recommendations, Assessment, Development and Evaluation \(GRADE\) gerada pelo software GRADEpro. \(ver anexo em PDF\)](#)

[Tabela III - Exemplo da tabela de Sumário de Resultados \(Summary of Findings \) do Sistema Grading of Recommendations, Assessment, Development and Evaluation \(GRADE\), gerada pelo software GRADEpro. \(ver anexo em PDF\)](#)

A sugestão do sistema GRADE é que sejam apresentados apenas 7 desfechos nas tabelas. Além disso, é possível observar que ambas as tabelas disponibilizam notas de rodapé com o motivo pelo qual o nível de evidência foi rebaixado ou aumentado, permitindo desta forma transparência no julgamento da evidência apresentada [28-29]. Essas tabelas podem ser geradas através de um site denominado GRADEpro GDT [33] e para as questões envolvendo testes diagnósticos há uma adaptação.

Recomendação

A força da Recomendação, preconizada pelo GRADE, reflete a extensão da nossa confiabilidade de que a estimativa do efeito está adequada para suportar uma decisão particular ou recomendação. É utilizada nas diretrizes e considera o equilíbrio entre os benefícios e danos de uma intervenção, variabilidade nos valores e preferências do paciente e o uso dos recursos (custos) da intervenção [4,6].

O GRADE apresenta apenas duas categorias de recomendação: ‘Forte’ ou ‘Fraca’. A Recomendação é forte quando os efeitos desejáveis da intervenção claramente superam os efeitos indesejáveis ou claramente não superam. Uma Recomendação Fraca resulta de resultados incertos, devido à qualidade da evidência ou quando os efeitos desejáveis e indesejáveis da intervenção estão balanceados [4,6].

Além disso, o GRADE considera que a força da recomendação também pode ser influenciada pelos valores e preferências do paciente e pelo uso dos recursos da intervenção. Quanto mais variável for a escolha do paciente quanto à intervenção proposta, maior a probabilidade de uma Fraca Recomendação. Neste espectro consideram-se as perspectivas, expectativas, crenças e objetivos do paciente em relação à saúde e à vida. Da mesma forma um custo mais alto enfraquecerá a recomendação [4,6] (Figura 2). Os custos aqui apresentados transcendem os custos monetários e são consideradas também outras unidades de custos diretos e indiretos como, por exemplo, dias de hospitalização, tempo e duração da terapia.



Figura 2 - Determinantes da Força de Recomendação pelo sistema Grading of Recommendations, Assessment, Development and Evaluation (GRADE).

O GRADE também prevê recomendações para o uso apenas no contexto da pesquisa quando há intervenções promissoras com prejuízos ou grandes custos detectados, mas com evidência insuficiente. O objetivo desta proposição é encorajar a pesquisa para responder questões clínicas ainda não elucidadas [34].

É importante apontar que o sistema GRADE não proporciona concordância na interpretação dos resultados. Embora todo o julgamento seja transparente e sistemático, não elimina discordâncias na interpretação desse julgamento. Não é objetivo do GRADE determinar uma avaliação objetiva da qualidade da evidência, processo, este, subjetivo e julgamento-dependente. O mérito desse sistema é permitir uma estrutura clara para guiar a interpretação dos pontos críticos julgados [17].

Implicações para a prática fisioterapêutica

O crescente uso de revisões sistemáticas e diretrizes para a prática clínica tem permitido ao fisioterapeuta embasar com evidência suas escolhas. Desta forma, com o intuito de aumentar a segurança na tomada de decisão, é importante que os fisioterapeutas entendam como o modelo vigente de avaliação da qualidade de evidência, em diretrizes e revisões sistemáticas, tem sido empregado e discutido.

A proposta do GRADE surge inserida em um modelo de raciocínio complexo, porém pragmático que ultrapassa a avaliação da evidência apenas nas considerações metodológicas intrínsecas dos estudos. O sistema aponta outros fatores que precisam ser considerados para uma análise global da evidência. Deixa de focar no risco de viés dos estudos, para averiguar também a variabilidade, a precisão e o direcionamento do resultado, além dos possíveis vieses provocados pela publicação. Em uma perspectiva filosófica permite um veredito que confronta o achado disponível e transparentemente julgado com um possível efeito verdadeiro. Permite avaliar quão confiante se está no efeito estimado, como reflexo da apreensão da realidade. Desconstrói, portanto, a tentativa dicotômica de classificar a evidência através de estudos de boa e má qualidade, além de desconsiderar uma relação direta entre nível de evidência e grau de recomendação, até então difundida por vários outros sistemas. Desta forma, proporciona uma tradução mais compatível com o mundo real e permite que o profissional de saúde transponha o conhecimento existente no seu espectro clínico de atuação.

Quando preconiza o foco em desfechos críticos e importantes para o paciente e considera seus valores e preferências, permite sobressair a imprescindível e irrestrita participação direta do paciente no paradigma da saúde baseada em evidência que orienta a inclusão dos contextos individuais na interpretação da evidência. Desta forma, o conhecimento da abordagem do sistema GRADE permite ao fisioterapeuta respaldar-se em escolhas clínicas informadas para que possa atuar no modelo de atuação prática centrada no paciente.

Conclusão

O sistema GRADE propõe uma metodologia compreensiva e sistemática para graduar e sumarizar a qualidade de evidência nas revisões sistemáticas e diretrizes e fornecer um suporte para o manejo das recomendações clínicas nas diretrizes. Classifica a qualidade da evidência em alta, moderada, baixa e muito baixa e a força da recomendação em forte e fraca. Através de julgamentos transparentes em cada etapa do processo permite o entendimento da evidência disponível fornecendo um suporte adequado para análise da confiança nos resultados encontrados. Está cada vez mais sendo incorporado nas revisões sistemáticas e diretrizes e, portanto, faz-se necessário o seu entendimento para facilitar a análise crítica da evidência e para que possa ser introduzido nas revisões sistemáticas em Fisioterapia.

Referências

1. GRADE working group. Grading quality of evidence and strength of recommendations. *BMJ* 2004;19:1490-94.
2. Atkins D, Eccles M, Flottorp S, Guyatt GH, Henry D, Hill 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. *BMC Health Serv Res* 2004;4(1):38.
3. BMJ Publishing Group Ltd. Article requirements. 2006. [citado 2016 Ago 7]. Disponível em URL: <http://resources.bmj.com/bmj/authors/article-submission/article-requirements>.
4. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coelho P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendation. *BMJ* 2008;336:924-26.
5. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HJ. What is "quality of evidence" and why is it important to clinicians? *BMJ* 2008;2008;336:995-8.
6. Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati, A et al. Going from evidence to recommendations. *BMJ* 2008;336:1049-51.
7. Schunemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al. Grading quality of evidence and strength of recommendation for diagnostic tests and strategies. *BMJ* 2008;336:1106-10.
8. Guyatt GH, Oxman AD, Kunz R, Jaeschke R, Helfand M, Liberati A et al. Incorporating considerations of resources use into grading recommendations. *BMJ* 2008;336:1170-3.
9. Jaeschke R, Guyatt GH, Dellinger P, Schunemann H, Levy MM, Kunz R et al. Use of GRADE grid to reach decisions on clinical practice guidelines when consensus is elusive. *BMJ* 2008;337.a744.
10. Guyatt GH, Oxman AD, Schunemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the *J Clin Epidemiol*. *J Clin Epidemiol* 2011;64:380-2.
11. Andrews JC, Schunemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Alonso-Coello P et al. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. *J Clin Epidemiol* 2013;66(7):726-35.
12. GRADE Working group. [citado 2016 Ago 10]. Disponível em URL: <http://www.gradeworkinggroup.org/>.
13. Davoli M, Amato L, Clark N, Farrell M, Hickman M, Hill S et al. The role of Cochrane reviews in informing international guidelines: a case study of using the Grading of Recommendation, Assessment, Development and Evaluation system to develop World Health Organization guidelines for the psychosocially assisted pharmacology treatment of opioid dependence. *Addiction* 2015;110(6):891-8.
14. Schünemann H, Brozek J, Guyatt G, Oxman A, eds. GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013. The GRADE Working Group, 2013. [citado 2016 Ago 10]. Disponível em: URL: www.guidelinedevelopment.org/handbook.
15. Harvey LA. GRADE the evidence. *J Physiother* 2013;59:5.
16. Austin TM, Richter RR, Sebelski CA. Introduction to the GRADE approach for guideline development: considerations for physical therapist practice. *Phys Ther* 2014;94:1652-9.
17. 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;64:401-6.
18. 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.

19. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 7. Rating the quality of evidence –inconsistency. *J Clin Epidemiol* 2011;64(12):1294-302.
20. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coelho P, Rind D. GRADE guidelines: 6. Rating the quality of evidence-imprecision. . *J Clin Epidemiol* 2011;64(12):1283-93.
21. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M et al. GRADE guidelines: 8. Rating the quality of evidence-indirectness. *J Clin Epidemiol* 2011;64(12):1303-10.
22. 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.
23. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Aki EA, Alonso-Coello P. GRADE guidelines: 9. Rating up the quality of evidence. *J Clin Epidemiol* 2011;64(12):1311-6.
24. Devereaux PJ, Choi PT, Lacchetti C, Weaver B, Schunemann HJ, Haines T et al. A systematic review and meta-analysis of studies comparing mortality rates of private for-profit and private not-for-profit hospitals. *CMAJ* 2002;166:1399-406.
25. Brozek JL, AKI EA, Compalati E, Kreis J, Terracciano L, Fiocchi A et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines Part 3 of 3. The GRADE approach to developing recommendations. *Allergy* 2011;66:588-95.
26. 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* 2013;64(4):395-40.
27. Guyatt G, Oxman AD, Sultan S, Brozek J, Glasziou P, Alonso-Coello P. 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):151-7.
28. GRADE guidelines: 1. Introduction. GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;383-94.
29. Guyatt GH, Oxman AD, Santesso N, Helfand M, Vist G, Kunz R. GRADE guidelines: 12. Preparing Summary of Findings tables-binary outcomes. *J Clin Epidemiol* 2013;66(2):158-72.
30. Guyatt GH, Thorlund K, Oxman AD, Walter SD, Patrick D, Furukawa TA et al. GRADE guidelines: 13. Preparing summary of findings tables and evidence profiles-continuous outcomes. *J Clin Epidemiol* 2013;66(2):173-83.
31. Kannan P, Claydon LS. Some physiotherapy treatments may relieve menstrual pain in women with primary dysmenorrhea: a systematic review. *J Physiother* 2014;60:13-21.
32. Zwerink M, Brusse-Keizer M, van der Valk PD, Zielhuis GA, Monnikhof EM, van der Palen J et al. Self management for patients with chronic obstructive pulmonary disease. *Cochrane Database Syst Rev* 2014;19(3): CD002990. doi: 10.1002/14651858.CD002990.pub3.
33. GRADEpro GDT. [citado 2016 Ago 10]. Disponível em URL: <https://grade.pro.org>.
34. Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *J Clin Epidemiol* 2013;66(7):719-25.