

4.5 Como distinguir entre evidências de alta e baixa qualidade

Nem todas as evidências são de alta qualidade e confiáveis para a tomada de decisão. Existem ferramentas para muitas (mas não todas) formas de evidências a fim de ajudar a julgar se as evidências (de um único estudo ou de um conjunto de evidências) podem ser confiáveis. Como descrevemos aqui, essas ferramentas usam pontuações ou graus para ajudar os usuários a entender quanta confiança podem ter na evidência. Muitos periódicos agora exigem que os autores sigam os padrões de relatórios, como CONSORT para ensaios clínicos randomizados e PRISMA para sínteses de evidências. A maioria dos periódicos não exige que os revisores usem ferramentas específicas para avaliar a qualidade dos estudos ou a força das recomendações. Como resultado, a publicação em um periódico revisado por pares não é um bom indicador de qualidade.

Problema	Resposta
Estudos (e diretrizes) variam em qualidade (ou confiabilidade)	<ul style="list-style-type: none"> Ferramentas de avaliação de qualidade (ou avaliação crítica) foram desenvolvidas para desenhos de estudo específicos (p. ex., ensaio clínico randomizado), para categorias abrangentes de desenhos de estudo (p. ex., estudo observacional, pesquisa qualitativa e síntese de evidências), e para diretrizes – consulte o anexo ao final deste capítulo (seção 4.16) para exemplos (RoB2, ROBINS-I, <i>checklist</i> do JBI, AMSTAR e AGREE II) As ferramentas podem produzir um resumo dos julgamentos (p. ex., baixo risco de viés usando RoB2 ou ROBINS-I), uma pontuação apresentada em intervalos (p. ex., alta qualidade usando AMSTAR), um conjunto de pontuações (p. ex., seis domínios usando AGREE II), ou um conjunto de considerações que podem informar um resumo dos julgamentos (p. ex., <i>checklist</i> do JBI)
Os conjuntos de evidências variam em sua certeza (ou na confiança que se pode ter neles)	<ul style="list-style-type: none"> Ferramentas para avaliar a certeza foram desenvolvidas para um conjunto de evidências abordando a mesma questão (p. ex., o efeito de uma intervenção sobre um desfecho específico ou o significado que os cidadãos atribuem a um fenômeno específico) – consulte a seção 4.16 para dois exemplos (GRADE e GRADE CERQual) As ferramentas podem produzir um resumo dos julgamentos sobre a confiança de que o efeito verdadeiro é similar ao efeito estimado (p. ex., alto grau de certeza com GRADE) ou que o fenômeno de interesse é bem representado por um achado de estudo qualitativo (com GRADE CERQual) Um resumo dos julgamentos sobre a certeza de uma estimativa de efeito é mais útil do que um teste de significância estatística demonstrando que uma intervenção “funciona” ou “não funciona” (o que acontecerá por acaso uma em cada 20 vezes se a significância estatística for definida no nível 0,05)
As recomendações variam quanto à sua força	<ul style="list-style-type: none"> Ferramentas de avaliação da força foram desenvolvidas para recomendações de diretrizes (p. ex., GRADE, além de classificar a certeza de um corpo de evidências, como descrito acima) – veja a seção 4.16 para um exemplo As ferramentas podem produzir um resumo dos julgamentos sobre se a maioria dos tomadores de decisão escolheria prosseguir com uma intervenção (p. ex., forte com GRADE) ou se a maioria precisaria avaliar cuidadosamente os prós e contras de uma intervenção
Algumas fontes de (ou abordagens usadas para gerar) evidências podem ser difíceis de julgar	<ul style="list-style-type: none"> Não existem ferramentas amplamente aceitas para avaliar o grau de confiança em: <ul style="list-style-type: none"> Um <i>expert</i>, embora existam exemplos como o <i>Good Judgment Project</i> (Projeto Bom Julgamento) para previsão (retomamos sobre a opinião de <i>experts</i> depois neste capítulo e, no caso da opinião de <i>experts</i> sobre parâmetros de modelo, na seção 4.16) Modelos usados na geração de algumas formas de evidência (que abordamos na seção 4.7, ao falar sobre modelos de mudanças climáticas, e na seção 4.16) Um algoritmo de inteligência artificial usado na geração de alguns tipos de evidências, embora exemplos como TRIPOD estejam começando a surgir (3)

Distinguir entre evidências de alta e baixa qualidade é particularmente desafiador quando as evidências estão incorporadas em painéis, modelos e outros formatos, e quando há conflitos de interesse em jogo. Retomamos sobre conflitos de interesse nas [seções 4.12, 4.14 e 4.16](#). Embora não seja o foco deste relatório, distinguir entre “dados brutos” de alta e baixa qualidade também pode ser um desafio, e organizações como o UNICEF desenvolveram estruturas de qualidade de dados para ajudar nisso (bit.ly/3DQQRrv).

Alguns “balcões únicos”, como o *Social Systems Evidence* e o inventário da Rede de Evidências para Apoiar a Tomada de Decisões na COVID-19 (COVID-END) (descrita na [seção 4.6](#)), usam algumas dessas ferramentas para que os tomadores de decisão e quem os apoia possam focar nas sínteses de evidências de alta qualidade ou entender que estão usando as melhores (se não forem de alta qualidade) sínteses de evidências disponíveis.

A pandemia de COVID-19 exigiu que os tomadores de decisão tomassem decisões difíceis em curtos períodos de tempo, inicialmente com poucas evidências, geralmente indiretas, e, com o tempo, com estudos, conjuntos de evidências e recomendações desenvolvidos por meio de um processo robusto. Para apoiar a tomada de decisão sobre a COVID-19 com base em conjuntos de evidências (em vez de estudos únicos), em seu inventário de sínteses das “melhores” evidências, a COVID-END traçou um perfil das sínteses que eram atualizadas (com base na data de busca por evidências) e de alta qualidade (com base na ferramenta AMSTAR), e avaliou a certeza das evidências (com base na ferramenta GRADE).

Assim como nem todas as evidências são de alta qualidade, nem todas as evidências globais serão aplicáveis em um determinado contexto. Por exemplo, uma síntese de evidências contendo estudos realizados somente em países de alta renda pode ter aplicabilidade limitada a alguns países de baixa renda. Pode haver diferenças importantes nas condições de base, nas realidades e restrições locais, e nos componentes estruturais do sistema local (p. ex., sistema nacional de saúde ou sistema educacional provincial/estadual). Uma ferramenta SUPPORT também pode ajudar as pessoas a pensar sobre a aplicabilidade local dos achados de uma síntese de evidências e considerar como ainda é possível extrair *insights*, mesmo quando os achados não são aplicáveis.(4)

O raciocínio Bayesiano tem atraído cada vez mais atenção como uma forma de redesenhar deliberadamente nossos “mapas mentais” sobre desafios e formas de respondê-los, não substituindo tudo que pensávamos saber por novas informações, mas modificando nosso entendimento para um grau adequado. Tal grau depende de quanta confiança se tinha em seu conhecimento anterior (a probabilidade “anterior” de algo ser verdadeiro) e quanta confiança se deposita no novo conhecimento. Mais confiança pode ser depositada no novo conhecimento se vier de uma síntese de evidências de alta qualidade que inclua estudos conduzidos em contextos semelhantes ao seu.



Intermediária e produtora de evidências, Gillian Leng

Executiva experiente, liderando uma agência de avaliação de tecnologias e diretrizes para apoiar a tomada de decisão no âmbito da atenção à saúde e assistência social para governos, provedores de serviços e pacientes

O Reino Unido tem liderado um trabalho ao longo de muitos anos que visa encorajar a síntese e o uso de evidências – desde o primeiro ensaio clínico randomizado para prevenir o escorbuto em marinheiros, até os mais recentes e inovadores *What Works Centres* para promover o uso de evidências em uma série de áreas políticas. Como parte desse movimento baseado em evidências, nos últimos 20 anos, o *National Institute for Health and Care Excellence* - NICE (Instituto Nacional para a Saúde e Excelência Clínica) transformou o uso de evidências em prática da atenção à saúde, bem como em iniciativas mais amplas de saúde pública e assistência social.

A pandemia de COVID-19 reforçou dramaticamente a necessidade de evidências de alta qualidade para informar políticas e práticas, e também destacou as consequências negativas da mídia social e da desinformação a ela associada. Nesse contexto, o trabalho da Comissão Global de Evidências para Responder aos Desafios Sociais é extremamente importante e deve ser visto como uma leitura essencial para todos os formuladores de políticas em todo o mundo.