

NÍVEIS DE EVIDÊNCIAS CIENTÍFICAS NA PRÁTICA MÉDICA

Regina El Dib

Para cada questão clínica na área da Saúde, existe um determinado desenho de estudo mais apropriado para responder esta questão, seja no sentido de boa validade interna (metodologia do desenho e condução do estudo), seja na boa validade externa (aplicabilidade dos resultados do estudo na prática clínica). Se quisermos saber sobre a evolução de uma determinada doença ou o prognóstico da mesma, o desenho de estudo mais apropriado para responder esta questão é um estudo coorte. Se quisermos saber sobre a sensibilidade e especificidade de um teste diagnóstico, o tipo de estudo que possui uma validade interna mais adequada é o estudo de acurácia. Todavia, se a pergunta é sobre tratamento ou profilaxia, procuramos por ensaios clínicos randomizados.

Este raciocínio fundamentado por diferentes desenhos de estudos vem tomando parte do cotidiano de vários profissionais da área da Saúde.

Até a alguns anos atrás, a prática clínica era embasada apenas por estudos fisiopatológicos, opiniões de especialistas no assunto, livros-textos, pesquisa *in vitro* e animal. Entretanto, a Medicina Baseada em Evidências (MBE) mudou este cenário facilitando o ensino e a pesquisa científica. Através de rigorosos métodos científicos, busca exaustiva pela literatura associada à experiência do clínico, tomamos decisões baseadas em evidências em nossa prática clínica.

Este processo originou-se com a Medicina Baseada em Evidências, sendo o “elo entre a boa ciência e a boa prática clínica” ⁽¹⁾. Deste modo, com pesquisas conduzidas com validade interna adequada, sem ocorrência de vieses (isto é, erro sistemático durante a condução, elaboração e planejamento de uma pesquisa científica), reduzimos a incerteza na prática clínica e aplicamos os resultados destas pesquisas em nossos pacientes com maior confiabilidade.

O termo MBE foi utilizado pela primeira vez na literatura médica em 1992, por pesquisadores da Universidade McMaster, sendo definida como o uso consciencioso, explícito e judicioso da melhor evidência disponível de pesquisa de cuidado médico na administração aos pacientes ^(2,3).

Frequentemente o clínico da área da Saúde quando diante de uma situação inesperada em seu consultório, hospital ou clínica, reconhece o próprio limite de fornecer uma

resposta de prontidão ao seu paciente. Por exemplo, imagine um médico atendendo um paciente com aneurisma da aorta abdominal (AAA), sendo que alguns aneurismas da aorta abdominal (AAAs) apresentam-se como emergências e necessitam de cirurgia e outros permanecem assintomáticos, e sabendo que o tratamento deste último depende de vários fatores, incluindo o tamanho do aneurisma. AAAs assintomáticos, com diâmetro > 5.5 cm, são geralmente operados enquanto que pequenos AAAs (< 4.0 cm em diâmetro) são monitorados com ultrassonografia⁽⁴⁾. Este paciente pergunta ao profissional se o reparo cirúrgico imediato é mais eficaz e seguro quando comparado a rotina de vigilância da ultrassonografia na redução de mortalidade. O profissional não consegue uma resposta de prontidão e reconhece sua incapacidade de fornecer uma informação precisa. Uma solução comum neste caso é recorrer a um especialista na área ou um livro-texto consagrado na literatura. Esta atitude revela poucas habilidades e conhecimento do profissional para lançar mão das ferramentas oferecidas pela MBE^(5,6). Imaginamos, agora, o mesmo profissional com conhecimentos e entendimento do processo de MBE. O médico volta para seu consultório, realiza uma estratégia de busca sensível para recuperar artigos sobre pacientes com AAA assintomático (com diâmetro entre 4.0 e 5.5 cm) submetidos ao reparo cirúrgico versus rotina de vigilância da ultrassonografia, acessa as principais bases de dados eletrônicas da área da Saúde, seleciona os títulos potenciais e realiza uma avaliação crítica destes artigos. Neste caso, se ele encontrar uma revisão sistemática (nível 1++ de evidências, de acordo com o *Scottish Intercollegiate Guidelines Network - SIGN*⁽⁷⁾ e nível 1a de acordo com o *Oxford Centre for Evidence-based Medicine Levels of Evidence - Oxford CEBM*⁽⁸⁾) com orientações específicas para a prática clínica, confirmando ou não a eficácia da cirurgia reparadora imediata na diminuição de mortalidade em pacientes com AAA assintomático, os resultados irão satisfazer o profissional e torná-lo apto a responder ao paciente a sua dúvida. Caso o profissional não encontrar estudos considerados sendo nível 1++⁽⁷⁾ ou 1a⁽⁸⁾ de evidências, o mais sensato é lidar com as provas científicas disponíveis no momento, estando ciente de seu nível de evidência e grau de recomendação.

Existem várias classificações quanto ao nível e grau de evidências na literatura. Apresentaremos neste capítulo o *Scottish Intercollegiate Guidelines Network – SIGN*⁽⁷⁾ e o *Oxford Centre for Evidence-based Medicine Levels of Evidence - Oxford CEBM*⁽⁸⁾,

El Dib R. Níveis de Evidências Científicas na Prática Médica (chapter 1). In: Guia Prático de Ultrassonografia Vascular 2º edição. 2011 Di Livros Editora Ltda., RJ, Brasil, ISBN: 978-85-86703-91-1.

duas ferramentas apropriadas e completas em relação aos níveis de evidências devido a suas considerações referente ao desenho e qualidade metodológica dos estudos (Apêndices 1, e 2). Quanto ao grau de evidências, utilizaremos um sistema simples, baseado no Oxford CEBM ⁽⁸⁾, e por nós modificado, que possibilita qualquer profissional avaliar a robustez dos artigos científicos relacionados com as intervenções terapêuticas (Apêndice 3).

No exemplo citado anteriormente, realizamos uma estratégia de busca com todos os descritores referentes ao termo AAA assintomático e relacionamos com o termo cirurgia e seus sinônimos – a partir da busca nas bases de dados Descritores em Ciências da Saúde (DECS, www.decs.bvs.br) e Medical Subject Headings (MeSH, www.pubmed.com). Após este procedimento, ‘rodamos’ a estratégia na base de dados *the Cochrane Library* – biblioteca virtual que contem inúmeras revisões sistemáticas e ensaios clínicos randomizados (ECRs) na área da saúde, sendo considerada a melhor fonte de evidências para a tomada de decisões nos cuidados à saúde (encontra-se disponível gratuitamente no site www.centrocochranedobrasil.org). O intuito de rodar a estratégia de busca primeiro na Biblioteca Cochrane fornece ao profissional mapear rapidamente os estudos existentes com alto nível de evidências (ECRs e revisões sistemáticas) e disponíveis naquele momento. Com os títulos identificados na Biblioteca Cochrane, selecionamos os artigos em potencial para uma leitura minuciosa. Caso não existam evidências sobre o determinado assunto, o profissional poderá lançar mão das principais bases de dados na área da saúde como a PUBMED e EMBASE para identificar estudos sobre o tópico de interesse.

É esse um dos objetivos da Medicina Baseada em Evidências, capacitar o profissional a desenvolver habilidades para melhorar a qualidade de seu atendimento. Além disso, a MBE tem como objetivos avaliar a literatura na área da Saúde, reduzir a incerteza clínica, reduzir probabilidades de erros, diminuir o autoritarismo no ensino, racionalizar custos e tornar a relação terapeuta ou médico/ paciente mais prazerosa ⁽⁵⁾.

A boa formulação de uma questão clínica é o alicerce de uma boa pesquisa. Para elaborarmos uma pergunta adequada sobre alguma questão a ser resolvida, devemos ter em mente quatro itens-chave, são eles: situação clínica (a doença), intervenção, grupo controle e desfecho clínico. Estes itens-chave formam a sigla PICO que em inglês

El Dib R. Níveis de Evidências Científicas na Prática Médica (chapter 1). In: Guia Prático de Ultrassonografia Vascular 2º edição. 2011 Di Livros Editora Ltda., RJ, Brasil, ISBN: 978-85-86703-91-1.

significa: *patient, intervention, control group* e *outcome*, sendo conhecida entre os revisores e pesquisadores internacionalmente.

Por exemplo, se quisermos saber qual frequência de dose do dinitrato de isossorbida é mais efetiva para a prevenção do desenvolvimento da tolerância ao nitrato em pacientes com doença das coronárias por meio da ecografia vascular e teste de esforço⁽⁹⁾, já temos uma pergunta completa: situação clínica – doença das coronárias; intervenção de interesse – dinitrato de isossorbida a cada 6 horas; grupo controle – dinitrato de isossorbida a cada 12 horas; e o desfecho primário de interesse seria nível de tolerância do nitrato medido pela ecografia vascular e teste de esforço. Pronto. Já partimos do ponto-chave para a realização de um projeto de pesquisa adequado, minimizando as chances de erros nos resultados, pois com uma pergunta claramente feita, dificilmente iremos nos perder durante o processo de pesquisa.

A Colaboração Cochrane enfatiza em sua metodologia elaborarmos a questão clínica da seguinte ordem: [intervenção] vs [grupo controle] para [situação clínica]⁽¹⁰⁾. Às vezes, em uma pergunta não há necessidade de inserir o desfecho de interesse, porém em todo o projeto de pesquisa este item deverá estar claramente descrito.

Tomamos outro exemplo para elucidar este tópico. Imaginem que a técnica de dispositivos de compressão pneumática intermitente (CPI) vem sendo explorada através de estudos para a comprovação desta abordagem em pacientes com úlcera varicosa⁽¹¹⁾. Como ficaria a pergunta clínica? Iniciamos pelos quatro itens-chave: situação clínica – úlcera varicosa; intervenção de interesse – CPI; grupo controle – nenhuma intervenção ou procedimento *sham* de CPI; e desfecho de interesse – aumento da cicatrização das feridas. Seguindo a estrutura da formulação de pergunta pela Colaboração Cochrane, temos: O CPI é mais efetivo e seguro quando comparado a nenhuma intervenção no aumento da cicatrização das feridas em pacientes com úlcera varicosa?

O objetivo e o título são uma derivação da pergunta. Neste último caso citado, o objetivo ficaria assim: Avaliar a eficácia e segurança dos dispositivos de compressão pneumática intermitente quando comparado a nenhuma intervenção no aumento da cicatrização das feridas em pacientes diagnosticados com úlcera varicosa. E por fim, o título poderia ser: Dispositivos de compressão pneumática intermitente versus nenhuma intervenção no tratamento de úlcera varicosa.

Não é difícil desenvolver uma questão clínica, basta termos os quatro itens-chave para resolver este problema. É claro que dependendo do desenho de estudo, os itens-chave poderão variar, mas sempre o desfecho estará presente. Em estudos coortes temos a situação clínica, o fator de risco e o desfecho de interesse. Por exemplo, variz associada à tromboflebite superficial pode ser considerada um fator de risco para o desenvolvimento de trombose venosa profunda? Neste caso, tromboflebite superficial é a situação clínica, variz, o fator de risco, e trombose venosa profunda o desfecho de interesse.

Efetividade, eficácia, eficiência e segurança

Quando abordamos tratamento e falamos em evidências, nos referimos à efetividade, eficiência, eficácia e segurança. Ao mencionarmos efetividade nos referimos ao tratamento que funciona em condições de mundo real. Eficácia, quando o tratamento funciona em condições de mundo ideal. Eficiência, quando o tratamento é barato e acessível para que os consumidores possam usufruí-lo. E por último, referimo-nos à segurança quando uma intervenção possui características confiáveis que inviabilizam causar algum efeito indesejável para o paciente.

Quando o propósito é saber se medicamentos funcionam em pacientes na UTI monitorados 24hs, a palavra mais sensata aqui é ‘eficácia’, pois o paciente receberá toda a medicação e procedimentos nos horários pré-estabelecidos pela equipe médica. Por outro lado, se o médico prescreve ao paciente uma medicação que deve ser tomada de oito em oito horas, por exemplo, e pede ao paciente que retorne ao consultório em duas semanas, o melhor termo aplicado para esta situação é ‘efetividade’, pois as condições são reais, variando desde esquecimento por parte do paciente de tomar o medicamento prescrito ou voltar ao consultório médico até falta de dinheiro para a compra do mesmo. A utilização correta destes termos auxilia no propósito do estudo, bem como, na aplicação correta dos resultados da pesquisa científica no cenário mais apropriado ao paciente.

Ensaio clínico randomizado e nível de evidências

Para responder questões sobre tratamento e prevenção, o melhor desenho de estudos são os ensaios clínicos randomizados (ECRs), considerados uma ferramenta poderosa e revolucionária da pesquisa científica, pois: (a) possuem pelo menos um grupo controle; (b) são prospectivos; (c) possuem processo de randomização (sorteio dos

participantes para serem alocados em um dos grupos de tratamento, possibilitando a todos os indivíduos a mesma chance de serem alocados tanto no grupo experimental como no grupo controle e, desta forma, distribuindo de forma homogênea as características dos participantes como sexo, idade, estágio de doença, peculiaridades genéticas, etc); (c) mascaramento dos desfechos a serem avaliados (com intuito de evitar viés/tendenciosidade por parte do avaliador e do paciente) ⁽¹²⁾. É considerado um estudo cujos participantes são alocados aleatoriamente para receber uma das várias intervenções clínicas. Os ensaios clínicos randomizados são experimentais, porque os investigadores podem influenciar no número e no tipo das intervenções, bem como na rotina, duração e dose de uma intervenção. Além disso, os ensaios clínicos randomizados são considerados quantitativos, comparativos e controlados ⁽¹²⁾.

Neste desenho de estudo, existem no mínimo dois grupos, um recebe a intervenção a ser testada (por exemplo, compressão com monitorização pelo eco-Doppler em cores para tratar de falso-aneurismas) e o outro grupo recebe outra intervenção, nenhuma intervenção ou placebo. Os dois grupos são seguidos de forma que os participantes não sejam perdidos até que os desfechos de interesse ocorram. O primeiro ensaio clínico publicado ocorreu em 1948 sobre o uso de estreptomicina para tuberculose.

Quando nos questionamos sobre efetividade de tratamentos e prevenção, a forma geralmente adequada é a exclusão de estudos não experimentais, ou seja, descartarmos estudos observacionais, uma vez que seus resultados podem fornecer conclusões positivas falsas sobre efetividade, eficácia e segurança. Devido à baixa ocorrência de vieses, os ensaios clínicos randomizados, especialmente as revisões sistemáticas, possuem uma metodologia adequada para nos instruir e, conseqüentemente, menor probabilidade de nos confundir em relação às questões terapêuticas. Por esse motivo, as revisões sistemáticas se tornaram o *gold standard* para julgarmos quando um tratamento é mais benéfico ou maléfico, quando comparado ao grupo controle.

Outrossim, o ECR é considerado nível 1++ de evidências de acordo com o *SIGN*⁽⁷⁾ e 1b baseado no Oxford CEBM⁽⁸⁾, mesmo quando há um risco de viés baixo, em outras palavras, quando, por exemplo, houver perdas e desistências menores que 20% do total do tamanho amostral em cada grupo durante o seguimento do estudo.

Para elucidar a questão de nível e grau de evidências, supomos querer saber sobre a efetividade da conversão do sirolimo versus redução do inibidor de calcineurina em

câncer de próstata em transplantados renais. Após a realização da estratégia de busca, identificação de artigos em potenciais e avaliação crítica dos mesmos, encontramos um ensaio clínico randomizado⁽¹³⁾: Nível de evidência 1++ (de acordo com o *SIGN*⁽⁷⁾) e 1b (baseado no Oxford CEBM⁽⁸⁾); grau de evidência A. Nós classificamos o grau de recomendação como A quando há evidências suficientemente fortes para recomendar a intervenção (tratamento benéfico).

Se por ventura, não encontrássemos ECR ou, então, nenhuma revisão sistemática, poderíamos basear nossa prática clínica com o próximo nível de evidências como, por exemplo, estudos coortes com boa qualidade metodológica (nível de evidência 2++, de acordo com o *SIGN*⁽⁷⁾ e nível 2b, baseado no Oxford CEBM⁽⁸⁾). Importante ressaltar que com um grau de recomendação A ou C (evidências suficientemente fortes para indicar ou contra-indicar uma intervenção, respectivamente) devemos sempre estar cientes do nível de evidência que estamos lidando e que as recomendações poderão ser modificadas conforme mais estudos no assunto estiverem disponíveis, sejam estes publicados ou não.

O grau de recomendação é classificado como B, quando temos evidências, mas que não são definitivas, ou seja, as evidências não são fortes o bastante para recomendar ou refutar uma intervenção. Ou então, quando não há nenhuma prova científica existente no momento sobre o determinado assunto.

Caso não possamos esperar pela finalização dos estudos, devemos seguir a hierarquia dos níveis de evidências e trabalhar com aquelas disponíveis no momento⁽¹⁴⁾.

Em suma, existem quatro passos para o profissional da área da Saúde praticar a MBE em qualquer situação quando surgir dúvidas sobre aspectos em relação ao diagnóstico clínico, tratamento, prevenção e prognóstico:

- Formular uma questão clínica clara baseada nos quatro itens-chave decorrente de um problema gerado pelo paciente;
- Pesquisar na literatura por relevantes artigos clínicos;
- Avaliar criticamente a evidência em relação à sua validade interna (metodologia) e aplicabilidade na prática clínica (validade externa);
- Implementar os achados úteis na prática clínica, considerando o nível e grau de evidências.

El Dib R. Níveis de Evidências Científicas na Prática Médica (chapter 1). In: Guia Prático de Ultrassonografia Vascular 2º edição. 2011 Di Livros Editora Ltda., RJ, Brasil, ISBN: 978-85-86703-91-1.

Neste último item, o médico pode tanto implementar diretamente as evidências em seus pacientes quanto utilizá-las para gerar diretrizes clínicas (*guidelines*).

Vale ressaltar que para realizar uma avaliação crítica da literatura, o profissional deve ter acesso aos diferentes métodos requeridos aos vários desenhos de estudo existentes e que foram desenvolvidos por diversos autores no Reino Unido e na América do Norte com intuito de capacitar indivíduos sem perícia no assunto a avaliar artigos clínicos. Estas estratégias podem ser utilizadas para avaliar artigos originais sobre diagnóstico, tratamento, prevenção e prognóstico, bem como avaliar revisões e meta-análises. Os sites www.sign.ac.uk e www.cebm.net apresentam diversos formulários de extração de dados e qualidade metodológica para os diferentes desenhos de estudos.

Referências

1. Castro AA. A pergunta da pesquisa. In: Atallah AN, Castro AA, editores. Medicina baseada em evidências: fundamentos da pesquisa clínica. São Paulo. Lemos editorial; 1998. p. 22-8.
2. Browman GP. Essence of Evidence-based medicine: a case report. Journal of Clinical Oncology 1999; 17:1999-1969.
3. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ 1996; 312(7023):71-2.
4. Ballard DJ, Filardo G, Fowkes G, Powell JT. Surgery for small asymptomatic abdominal aortic aneurysms. Cochrane Database Syst Rev. 2008;(4):CD001835.
5. Drummond JP, Silva E, Coutinho M. Medicina baseada em evidências: novo paradigma assistencial e pedagógico. 2 ed. São Paulo: Editora Atheneu, 2004.
6. Gomes MM. Medicina baseada em evidências princípios e práticas. 2 ed. São Paulo:Reichmann & Autores Editores, 2006.
7. Scottish Intercollegiate Guidelines Network (SIGN). Clinical guidelines: criteria for appraisal for national use. Edinburgh: SIGN 50; 2004.
8. Centre for Evidence Based Medicine (CEBM). Levels of evidence. Available in URL at <http://www.cebm.net/index.aspx?o=1025>
9. Yu GL, Zhuang HP, Wang ZH, Liao JT. Evaluation of nitrate tolerance in patients with coronary heart disease by vascular ultrasonography and treadmill exercise. Int J Cardiol 1999; 69(2):133-7.
10. Cochrane Reviewers' Handbook 4.2.2. Centro Cochrane do Brasil. Disponível em URL: <http://www.centrocochranedobrasil.org>.
11. Nelson EA, Mani R, Vowden K. Intermittent pneumatic compression for treating venous leg ulcers. Cochrane Database Syst Rev. 2008; (2):CD001899.
12. Jadad A. Randomised controlled trials: a user's guide. BMJ 1998; 1:1-3.
13. Watson CJ, Gimson AE, Alexander GJ, Allison ME, Gibbs P, Smith JC, et al. A randomized controlled trial of late conversion from calcineurin inhibitor (CNI)-based to sirolimus-based immunosuppression in liver transplant recipients with impaired renal function. Liver Transpl 2007; 13(12):1694-702.
14. Sackett DL, Rosenberg WMC, Gray JAM, et al: Evidence-based medicine: What it is and what it isn't. BMJ 1996; 312:71-2.

APÊNDICE 1

LEVEL OF EVIDENCE ⁽⁷⁾

1++ High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias

1+ Well conducted meta-analyses, systematic reviews of RCTs, or RCTs with a low risk of bias

1 Meta-analyses, systematic reviews or RCTs, or RCTs with a high risk of bias

2++ High quality systematic reviews of case-control or cohort studies or high quality case-control or cohort studies with a very low risk of confounding, bias, or chance and a high probability that the relationship is causal

2+ Well conducted case-control or cohort studies with a low risk of confounding, bias, or chance and a moderate probability that the relationship is causal

2 Case-control or cohort studies with a high risk of confounding, bias, or chance and a significant risk that the relationship is not causal

3 Non-analytic studies, eg. case reports, case series

4 Expert opinion

APÊNDICE 2

Oxford Centre for Evidence-based Medicine Levels of Evidence (May 2001)⁽⁸⁾

Level	Therapy/Prevention, Aetiology/Harm	Prognosis	Diagnosis	Differential diagnosis/symptom prevalence study	Economic and decision analyses
1a	SR (with <u>homogeneity*</u>) of RCTs	SR (with <u>homogeneity*</u>) of inception cohort studies; <u>CDR†</u> validated in different populations	SR (with homogeneity*) of Level 1 diagnostic studies; CDR† with 1b studies from different clinical centres	SR (with homogeneity*) of prospective cohort studies	SR (with homogeneity*) of Level 1 economic studies
1b	Individual RCT (with narrow <u>Confidence Interval‡</u>)	Individual inception cohort study with ≥ 80% follow-up; <u>CDR†</u> validated in a single population	Validating** cohort study with good††† reference standards; or CDR† tested within one clinical centre	Prospective cohort study with good follow-up****	Analysis based on clinically sensible costs or alternatives; systematic review(s) of the evidence; and including multi-way sensitivity analyses
1c	<u>All or none§</u>	All or none case-series	Absolute SpPins and SnNouts††	All or none case-series	Absolute better-value or worse-value analyses ††††
2a	SR (with <u>homogeneity*</u>) of cohort studies	SR (with <u>homogeneity*</u>) of either retrospective cohort studies or untreated control groups in RCTs	SR (with homogeneity*) of Level >2 diagnostic studies	SR (with homogeneity*) of 2b and better studies	SR (with homogeneity*) of Level >2 economic studies
2b	Individual cohort study (including low quality RCT; e.g., <80% follow-up)	Retrospective cohort study or follow-up of untreated control patients in an RCT; Derivation of <u>CDR†</u> or validated on split-sample§§§ only	Exploratory** cohort study with good††† reference standards; CDR† after derivation, or validated only on split-sample§§§ or databases	Retrospective cohort study, or poor follow-up	Analysis based on clinically sensible costs or alternatives; limited review(s) of the evidence, or single studies; and including multi-way sensitivity analyses
2c	"Outcomes" Research; Ecological studies	"Outcomes" Research		Ecological studies	Audit or outcomes research
3a	SR (with <u>homogeneity*</u>) of case-control studies		SR (with homogeneity*) of 3b and better studies	SR (with homogeneity*) of 3b and better studies	SR (with homogeneity*) of 3b and better studies

El Dib R. Níveis de Evidências Científicas na Prática Médica (chapter 1). In: Guia Prático de Ultrassonografia Vascular 2º edição. 2011 Di Livros Editora Ltda., RJ, Brasil, ISBN: 978-85-86703-91-1.

3b	Individual Case-Control Study		Non-consecutive study; or without consistently applied reference standards	Non-consecutive cohort study, or very limited population	Analysis based on limited alternatives or costs, poor quality estimates of data, but including sensitivity analyses incorporating clinically sensible variations.
4	Case-series (and <u>poor quality cohort and case-control studies</u>)	Case-series (and <u>poor quality prognostic cohort studies</u>)	Case-control study, poor or non-independent reference standard	Case-series or superseded reference standards	Analysis with no sensitivity analysis
5	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"	Expert opinion without explicit critical appraisal, or based on economic theory or "first principles"

Produced by Bob Phillips, Chris Ball, Dave Sackett, Doug Badenoch, Sharon Straus, Brian Haynes, Martin Dawes since November 1998.

Notes

Users can add a minus-sign "-" to denote the level of that fails to provide a conclusive answer because of:

- EITHER a single result with a wide Confidence Interval (such that, for example, an ARR in an RCT is not statistically significant but whose confidence intervals fail to exclude clinically important benefit or harm)
- OR a Systematic Review with troublesome (and statistically significant) heterogeneity.
- Such evidence is inconclusive, and therefore can only generate Grade D recommendations.

*	By homogeneity we mean a systematic review that is free of worrisome variations (heterogeneity) in the directions and degrees of results between individual studies. Not all systematic reviews with statistically significant heterogeneity need be worrisome, and not all worrisome heterogeneity need be statistically significant. As noted above, studies displaying worrisome heterogeneity should be tagged with a "-" at the end of their designated level.
†	Clinical Decision Rule. (These are algorithms or scoring systems which lead to a prognostic estimation or a diagnostic category.)
‡	See note #2 for advice on how to understand, rate and use trials or other studies with wide confidence intervals.
§	Met when <u>all</u> patients died before the Rx became available, but some now survive on it; or when some patients died before the Rx became available, but <u>none</u> now die on it.
§§	By poor quality <u>cohort</u> study we mean one that failed to clearly define comparison groups and/or failed to measure exposures and outcomes in the same (preferably blinded), objective way in both exposed and non-exposed individuals and/or failed to identify or appropriately control known confounders and/or failed to carry out a sufficiently long and complete follow-up of patients. By poor quality <u>case-control</u> study we mean one that failed to clearly define comparison groups and/or failed to measure exposures and outcomes in the same (preferably

El Dib R. Níveis de Evidências Científicas na Prática Médica (chapter 1). In: Guia Prático de Ultrassonografia Vascular 2º edição. 2011 Di Livros Editora Ltda., RJ, Brasil, ISBN: 978-85-86703-91-1.

	blinded), objective way in both cases and controls and/or failed to identify or appropriately control known confounders.
§§§	Split-sample validation is achieved by collecting all the information in a single tranche, then artificially dividing this into "derivation" and "validation" samples.
††	An "Absolute SpPin" is a diagnostic finding whose <u>Specificity</u> is so high that a <u>Positive</u> result rules- <u>in</u> the diagnosis. An "Absolute SnNout" is a diagnostic finding whose <u>Sensitivity</u> is so high that a <u>Negative</u> result rules- <u>out</u> the diagnosis.
‡‡	Good, better, bad and worse refer to the comparisons between treatments in terms of their clinical risks and benefits.
†††	<u>Good</u> reference standards are independent of the test, and applied blindly or objectively to applied to all patients. <u>Poor</u> reference standards are haphazardly applied, but still independent of the test. Use of a non-independent reference standard (where the 'test' is included in the 'reference', or where the 'testing' affects the 'reference') implies a level 4 study.
††††	Better-value treatments are clearly as good but cheaper, or better at the same or reduced cost. Worse-value treatments are as good and more expensive, or worse and the equally or more expensive.
**	Validating studies test the quality of a specific diagnostic test, based on prior evidence. An exploratory study collects information and trawls the data (e.g. using a regression analysis) to find which factors are 'significant'.
***	By poor quality prognostic cohort study we mean one in which sampling was biased in favour of patients who already had the target outcome, or the measurement of outcomes was accomplished in <80% of study patients, or outcomes were determined in an unblinded, non-objective way, or there was no correction for confounding factors.
****	Good follow-up in a differential diagnosis study is >80%, with adequate time for alternative diagnoses to emerge (eg 1-6 months acute, 1 - 5 years chronic)

El Dib R. Níveis de Evidências Científicas na Prática Médica (chapter 1). In: Guia Prático de Ultrassonografia Vascular 2º edição. 2011 Di Livros Editora Ltda., RJ, Brasil, ISBN: 978-85-86703-91-1.

APÊNDICE 3

Grades of guidelines recommendations, modified according to (Oxford Reference) ⁽⁸⁾

Grade	Nature of recommendation
A	Based on clinical studies of good quality and consistency addressing the specific recommendations and including at least one randomized trial
B	Inconclusive or inconsistent studies of any level
C	Evidence based on clinical studies sufficiently strong to refute the intervention