

Avaliação Crítica e Implementação Prática da Evidência

Luís Filipe Azevedo¹, Altamiro da Costa Pereira¹

RESUMO

A Medicina Baseada na Evidência (MBE) é genericamente definida como a aplicação consciente, explícita e criteriosa da melhor evidência científica disponível na tomada de decisões sobre o cuidado individual dos doentes. Nesta rubrica de Pediatria Baseada na Evidência será apresentado um conjunto de artigos abordando aspectos conceptuais, metodológicos e operacionais relativos à prática da MBE no âmbito específico da Pediatria. Neste artigo desenvolvem-se questões teóricas e práticas relacionadas com a adequada avaliação crítica e implementação prática da evidência científica. O objectivo deste artigo é sugerir uma metodologia prática e sistemática de avaliação crítica da evidência incluindo três questões essenciais: (1) avaliação da validade ou qualidade metodológica dos artigos; (2) avaliação da importância científica e prática dos seus resultados e (3) avaliação da aplicabilidade prática dos mesmos. São apresentados os principais tipos de questões e estudos mais frequentemente aplicados na prática clínica, são apresentadas as hierarquias de níveis de evidência e grau de recomendação, é discutido o conceito de erro sistemático ou viés e sua classificação e é, finalmente, apresentado um conjunto de critérios de avaliação de qualidade metodológica, importância e aplicabilidade da evidência científica.

Nascer e Crescer 2008; 17(1): 30-36

¹ Serviço de Bioestatística e Informática Médica da Faculdade de Medicina da Universidade do Porto e Centro de Investigação em Tecnologias e Sistemas de Informação em Saúde – CINTESIS.

INTRODUÇÃO

No primeiro artigo desta série^[1] foram abordados aspectos relacionados com a definição, enquadramento, evolução e importância da medicina baseada na evidência (MBE), em particular no respeitante à área da Pediatria. A MBE foi genericamente definida como a aplicação consciente, explícita e criteriosa da melhor evidência científica disponível na tomada de decisões sobre o cuidado individual dos doentes^[2-4]. Adicionalmente foi proposto um processo faseado que facilita a sua implementação na prática clínica diária e que é constituído por cinco fases^[3]:

- 1) Aproveitar as oportunidades que surgem diariamente e que apelam à adequada e oportuna formulação de questões clinicamente relevantes;
- 2) Realização de uma eficaz pesquisa da evidência científica de maior qualidade na resposta à questão formulada;
- 3) Realização de uma adequada análise crítica da evidência encontrada, relativamente à sua validade, impacto e aplicabilidade na prática clínica;
- 4) A integração da evidência científica encontrada com a experiência clínica individual, com o contexto específico da prática clínica e com as características individuais (biológicas, psicológicas e sociais) de cada doente;
- 5) A avaliação deste processo com o objectivo de maximizar a sua eficácia e eficiência, tornando-o mais facilmente aplicável na prática clínica diária, minimizando o tempo e esforço dispendidos e maximizando o ganho em termos de qualidade de serviços prestados e formação do próprio profissional de saúde.

No segundo artigo desta série^[5] foram desenvolvidos em maior detalhe os aspectos relacionados com a adequada e oportuna formulação de questões clínicas e a eficaz pesquisa da evidência científica mais relevante, com especial ênfase nos recursos disponíveis na área da Pediatria.

Após a identificação eficaz da evidência, é necessário que o clínico seja capaz de rapidamente avaliar as características metodológicas mais relevantes dos artigos, em função do tipo de questão analisada, com o objectivo de fazer uma adequada selecção da evidência científica mais relevante. Após a selecção da evidência é necessário avaliar de forma adequada e detalhada a qualidade científica dessa evidência e a sua aplicabilidade prática. Neste terceiro artigo da série serão abordadas questões relacionadas com a avaliação crítica da evidência e será proposta uma metodologia sistemática de avaliação crítica que inclui três questões essenciais:

- 1) avaliação da validade ou qualidade metodológica dos artigos;
- 2) avaliação da importância científica e prática dos seus resultados e
- 3) avaliação da aplicabilidade prática dos mesmos.

AVALIAÇÃO CRÍTICA DA EVIDÊNCIA

Na actualidade, felizmente, assiste-se a um crescimento exponencial do número de artigos científicos publicados anualmente em qualquer área do saber, e em particular na área biomédica, o que constitui um sinal positivo, indicativo de que há cada vez mais ciência a ser feita e divulgada. Esta realidade tem, no entanto, que ser devidamente enquadrada, pois existe uma enorme heterogeneidade quanto à relevância, impacto e adequa-

ção metodológica, sendo que uma parte importante daquilo que é publicado é pouco credível ou tem uma utilidade prática muito limitada. Quando o clínico procura evidência científica da maior qualidade torna-se clara a importância de uma adequada avaliação crítica da evidência quanto à sua qualidade metodológica e aplicabilidade.

Hoje, existem cada vez mais recursos e formatos que fazem a recolha e avaliação da evidência em vez do clínico, ou utilizador final, de forma a facilitar estas tarefas. São bons exemplos os estudos de síntese de evidência de uma forma geral (revisões sistemáticas, estudos de meta-análise, etc.) e vários serviços periódicos de selecção e avaliação de evidência existentes (ex: a revista da Academia Americana de Pediatria – *AAP Grand Rounds*, disponível no endereço <http://aapgrandrounds.aappublications.org>). No entanto, estes formatos e recursos, sendo alguns deles excelentes, não resolvem o problema e pelo contrário, são ainda muito limitados no seu âmbito, pelo que em última análise cabe ao clínico, ou utilizador final da evidência científica, a tarefa de a avaliar devidamente quanto à sua adequação e aplicabilidade.

Neste contexto, e dada a importância deste tipo de competências, o objectivo deste artigo é sugerir uma metodologia prática e sistemática de avaliação crítica da evidência incluindo três questões essenciais^[3,6]:

- 1) avaliação da validade ou qualidade metodológica dos artigos;
- 2) avaliação da importância científica e prática dos seus resultados e
- 3) avaliação da aplicabilidade prática dos mesmos.

A implementação de uma metodologia deste tipo implica a identificação e aplicação de um conjunto de critérios de avaliação para cada um desses aspectos e em função do tipo de questão científica analisada. Obviamente, os critérios a utilizar vão depender da metodologia aplicada e da natureza da questão científica a que cada artigo responde. Por exemplo, os critérios para avaliar um artigo sobre a eficácia de uma intervenção terapêutica serão, naturalmente, diferentes dos apli-

cados na avaliação de um artigo sobre diagnóstico ou prognóstico.

Felizmente, é possível identificar um conjunto nuclear e limitado de questões e tipos de estudos mais frequentemente aplicados na prática clínica, sendo possível definir para cada um deles um conjunto de critérios de avaliação de qualidade e aplicabilidade. Os tipos de questões e estudos mais frequentemente encontrados são:

- 1) questões sobre frequência de fenómenos, manifestações ou entidades clínicas;
- 2) questões sobre diagnóstico e propriedades de testes diagnósticos;
- 3) questões sobre intervenções terapêuticas ou preventivas;
- 4) questões sobre prognóstico;
- 5) questões sobre etiologia e risco e
- 6) os estudos de síntese de evidência.

Os leitores mais interessados poderão consultar várias listas estruturadas, mais completas, de itens para avaliação da qualidade de cada um destes tipos de estudos^[3,6].

Em seguida descrever-se-ão os aspectos mais relevantes relacionados com cada uma das fases propostas para o processo de avaliação crítica da evidência científica.

1. AVALIAÇÃO DA VALIDADE OU QUALIDADE METODOLÓGICA

O processo de avaliação crítica da evidência deverá começar por uma adequada avaliação do desenho do estudo com o objectivo de verificar se este responde adequadamente à questão científica colocada, avaliando portanto a qualidade metodológica do estudo e a validade da resposta dada.

Interessa a este respeito sublinhar que existem várias abordagens metodológicas ou tipos de estudos distintos para responder a uma mesma questão de investigação, e logicamente existem alguns tipos de estudos mais adequados do que outros. Veja-se, por exemplo, que para responder a uma questão sobre a eficácia de uma intervenção terapêutica poder-se-ão utilizar estudos observacionais, no entanto, um estudo experi-

mental, por exemplo, um ensaio clínico aleatorizado, permitirá dar uma resposta potencialmente mais válida e adequada a uma questão deste tipo. Neste contexto, interessa discutir as designadas hierarquias de níveis de evidência, hoje muito frequentemente utilizadas no processo de avaliação crítica de evidência científica^[7,8]. Estas hierarquias não são mais que listas, ordenadas numa estrutura hierárquica, de vários tipos de estudos, que podem responder a uma mesma questão de investigação, e que tentam distingui-los em função da qualidade metodológica ou potencial validade das suas respostas^[7,8]. Existem hoje muitas e distintas hierarquias de evidência disponíveis, no entanto todas elas seguem uma mesma filosofia e são bastante semelhantes^[7,9,10], recomenda-se, em particular, a hierarquia do *Oxford Centre for Evidence-based Medicine (OC-EBM)*, que é uma das mais frequentemente utilizadas (disponível no endereço <http://www.cebm.net>). No respeitante, por exemplo, aos estudos sobre a eficácia de intervenções terapêuticas são considerados, habitualmente, os seguintes níveis de evidência:

- Revisões sistemáticas e estudos de meta-análise de ensaios clínicos aleatorizados (nível 1a do *OC-EBM*);
- Ensaios clínicos aleatorizados (nível 1b do *OC-EBM*);
- Estudos de intervenção não aleatorizados (nível 2 e 3 do *OC-EBM*);
- Estudos observacionais (nível 2 e 3 do *OC-EBM*);
- Estudos não controlados (nível 4 do *OC-EBM*);
- Opinião de especialistas (nível 5 do *OC-EBM*).

Apesar da utilidade destas hierarquias, estas não são suficientes para uma adequada avaliação da qualidade da evidência, uma vez que existem para cada tipo de estudo um conjunto de critérios que devem ser respeitados e avaliados, de forma a maximizar a validade dos resultados, através da minimização de várias potenciais fontes de enviesamento.

O termo viés é um sinónimo de erro sistemático e refere-se, precisamente, às ameaças à validade dos resultados de

um estudo na dependência de este tipo de erros. Os erros sistemáticos são erros que afectam de uma forma previsível, mas infelizmente nem sempre conhecida, os resultados de um estudo. Obviamente, a existência de erros sistemáticos ou vieses num estudo podem motivar conclusões erradas, com todas as consequências que isto acarreta, em particular na área médica. Assim, pode-se afirmar que a avaliação da validade ou qualidade metodológica de um determinado estudo está na directa dependência da avaliação das estratégias implementadas com o objectivo de prevenir as várias e heterogéneas fontes potenciais de vieses.

Existe uma enorme variedade de vieses, na dependência do mecanismo ou da fonte a partir da qual surge o erro sistemático, havendo também grande heterogeneidade na forma de os sistematizar ou classificar⁽¹¹⁾. Classicamente, no entanto, distinguem-se três tipos principais de vieses: vieses de selecção, vieses de informação ou medição e vieses devidos ao confundimento.

Os vieses de selecção são erros sistemáticos associados aos métodos de selecção de participantes. Os vieses de informação ou medição são erros sistemáticos associados à classificação ou medição das variáveis dependentes (variáveis de resposta) ou independentes (variáveis preditivas, intervenção ou exposição). Os vieses devidos ao confundimento surgem quando existem variáveis de confusão que podem interferir na estimação do efeito pretendido.

Em seguida apresenta-se um sumário dos critérios de qualidade a aplicar para os tipos mais frequentes de questões de investigação neste âmbito. Estes critérios representam os aspectos mais relevantes a ter em atenção na avaliação da validade ou qualidade metodológica dos estudos e estão directamente associados à prevenção dos principais tipos de vieses atrás referidos. Está fora do âmbito deste artigo uma explicação detalhada de cada um dos critérios apresentados, no entanto, em números subsequentes desta série serão explorados, e explicados em maior detalhe, no contexto de exemplos práticos direccionados para os principais tipos de questões e estudos identificados.

Estudos sobre intervenções terapêuticas ou preventivas

Este tipo de estudos pretendem, geralmente, avaliar a eficácia e segurança de intervenções terapêuticas, logo são estudos que pretendem responder a questões que dizem respeito à existência de relações causais entre uma intervenção e um determinado resultado. Neste tipo de questões os desenhos mais robustos do ponto de vista metodológico são os estudos experimentais, por exemplo, na forma de ensaios clínicos aleatorizados. Neste contexto, definem-se habitualmente os seguintes critérios de qualidade mais relevantes^(3,12):

- Houve uma atribuição aleatória das intervenções (aleatorização ou randomização)?
- A lista de aleatorização foi ocultada aos profissionais responsáveis pela atribuição da intervenção (*allocation concealment*)?
- O seguimento (*follow-up*) dos doentes foi suficientemente longo e completo?
- A análise respeita a atribuição inicial das intervenções (análise segundo o princípio da intenção de tratar – *Intention-to-treat analysis*)?
- Existiu ocultamento (*blinding* ou *masking*) do grupo de tratamento em todos os parceiros envolvidos no estudo (doentes, médicos assistentes, investigadores responsáveis pela análise de variáveis de resultados, investigadores responsáveis pela análise dos dados – *single blind, double blind, triple blind*)?
- Ambos os grupos experimental e de controlo receberam igual cuidado e tratamento à excepção das intervenções em estudo?
- Os grupos experimental e de controlo eram comparáveis, no início do estudo, relativamente a outras variáveis que possam, eventualmente influenciar os resultados do mesmo?

Estudos sobre validade de testes diagnósticos

Este tipo de estudos pretendem, geralmente, avaliar a capacidade de testes diagnósticos na discriminação entre populações doentes e não doen-

tes, implicando, assim, habitualmente, a comparação entre os resultados de um teste diagnóstico e os resultados de um teste padrão (*gold standard*), aplicados concorrentemente a uma amostra apropriada de indivíduos. Neste contexto, definem-se habitualmente os seguintes critérios de qualidade mais relevantes^(3,13):

- O teste diagnóstico foi avaliado num espectro apropriado de doentes (semelhantes àqueles a quem o aplicaríamos na prática)?
- A selecção de participantes foi feita com base populacional? Ou foi feita uma amostragem independente de casos e de controlos?
- A selecção de participantes foi feita de forma consecutiva?
- Quais foram os critérios de selecção utilizados? Estes reflectem adequadamente o espectro de doença e não doença encontrados na prática?
- Qual é o período de referência do estudo? Prospectivo ou retrospectivo?
- O teste padrão de diagnóstico (*gold standard*) foi aplicado em todos os participantes e de forma independente do resultado do teste em estudo?
- Existe uma comparação cega (implementação de estratégias de ocultamento) e independente com o teste padrão de diagnóstico?
- Os métodos e protocolos de aplicação do teste em avaliação e do teste padrão são descritos de forma adequada?
- O teste (ou conjunto de testes) a avaliar foi validado num segundo grupo independente de doentes ou noutros centros?

Estudos sobre prognóstico

Os estudos sobre prognóstico têm como objectivo essencial descrever a evolução da doença e os factores eventualmente associados a um melhor ou pior prognóstico, habitualmente considerando vários tipos de variáveis de resultado clínico relevantes na determinação do mesmo (mortalidade, morbidade, incapacidade, qualidade de vida, etc.). A resposta a este tipo de questões é habitualmente encontrada na forma de estudos observacionais. Neste contexto, definem-se habitualmente os seguintes

critérios de qualidade mais relevantes^[3,14]:

- Os participantes do estudo constituem uma amostra representativa de doentes, recrutados num ponto comum (geralmente inicial – *inception cohort*) de evolução da sua doença ou condição clínica?
- O seguimento dos doentes foi suficientemente longo e completo?
- A análise das variáveis de resultado clínico (*outcome*) obedeceu a critérios objectivos e foi feita aplicando estratégias de ocultamento adequadas?
- Ao identificar subgrupos com prognósticos diferentes, portanto factores que afectam o prognóstico, a hipótese foi considerada a priori ou foi encontrada só na fase de análise dos dados?
- Houve ajustamento para factores prognósticos importantes e já conhecidos, eventualmente factores de confusão da(s) associação(ões) de interesse?
- Houve um estudo de validação dos resultados, factores ou modelos encontrados num grupo independente de doentes ou num centro diferente?

Estudos sobre etiologia ou risco

Os estudos sobre etiologia ou risco pretendem, habitualmente, descrever factores associados ao desenvolvimento de doença. Neste tipo de estudos, no plano metodológico, os desenhos mais frequentemente encontrados são estudos observacionais. Neste contexto, definem-se geralmente os seguintes critérios de qualidade mais relevantes^[3, 15]:

- Houve uma definição clara dos grupos que estão a ser estudados e comparados?
- Os grupos são comparáveis em todos os aspectos à excepção da exposição em estudo (no caso dos estudos de coorte ou ensaios clínicos aleatorizados)? Especialmente no âmbito dos estudos observacionais não é possível garantir este critério, no entanto, idealmente dever-se-á procurar garantir a comparabilidade dos grupos e/ou implementar estratégias, na fase de análise dos dados, que permitam corrigir este problema (estratificação, ajustamento estatístico de medidas de associação, modelação multivariada, etc.).

- Os grupos pertencem a uma mesma população de referência e tiveram iguais possibilidades de estarem sujeitos à exposição em causa (no caso dos estudos de casos e controlos)?
- As variáveis de exposição e/ou resultado clínico (*outcome*) foram medidas de forma igual nos dois grupos?
- A medição das variáveis em estudo foi feita de forma válida e precisa?
- A medição das variáveis em estudo foi feita aplicando estratégias de ocultamento adequadas (quando estas são possíveis)?
- O seguimento dos doentes foi suficientemente longo e completo (no caso dos estudos de coorte ou ensaios clínicos aleatorizados)?
- Caso o objectivo do estudo seja o teste de hipóteses sobre relações causais, os resultados ou associações encontradas no estudo verificam os critérios de causalidade comumente aceites (relação temporal, magnitude da associação, gradiente de dose-resposta, reversibilidade de efeito ou estudos de suspensão/reintrodução, consistência com os resultados de outros estudos e plausibilidade biológica)?

Estudos de síntese de evidência

Os estudos de síntese de evidência pretendem sumariar ou agregar os resultados de estudos primários relativos a uma determinada questão de investigação, avaliar a existência de heterogeneidade nesses resultados e estudar as fontes que, eventualmente, explicam essa mesma heterogeneidade. Neste contexto, definem-se habitualmente os seguintes critérios de qualidade mais relevantes^[3,16]:

- Que tipo de estudos foram incluídos?
- Tem um objectivo claramente definido?
- As características dos participantes, intervenções e variáveis de resultado (*outcome*) dos estudos primários seleccionados são adequadamente descritas?
- A secção “Métodos” deverá incluir os seguintes tópicos:
 - Métodos de pesquisa que garantam, o melhor possível, terem sido encontrados todos os estudos relevantes e evi-

tem os chamados viéses de publicação. Idealmente deverão ser utilizadas várias estratégias complementares de pesquisa bibliográfica – pesquisa em várias bases de dados bibliográficas (Medline; Embase; BD da Cochrane Collaboration); pesquisa manual em revistas científicas seleccionadas, livros de resumos de congressos e encontros; pesquisa em listas de referências dos artigos encontrados; consulta de especialistas na matéria; etc.

- Critérios de inclusão e exclusão dos estudos primários adequados e explicitamente apresentados.
- Descrição adequada do processo de selecção dos artigos e avaliação da reprodutibilidade deste processo.
- Descrição dos métodos de extracção de dados nos estudos primários.
- Descrição dos métodos de avaliação da qualidade/validade dos estudos primários.
- Descrição dos métodos estatísticos eventualmente aplicados na síntese quantitativa dos resultados dos estudos primários.
- Descrição dos métodos utilizados para avaliação da heterogeneidade e de viéses de publicação.
- A secção de resultados deverá conter os seguintes tópicos:
 - Descrição do processo de selecção dos estudos primários e reprodutibilidade deste processo.
 - Sumário qualitativo das características dos estudos primários.
 - Sumário qualitativo da avaliação da qualidade/validade dos estudos primários.
 - Sumário quantitativo dos resultados dos estudos primários (quando adequado).
 - Avaliação da heterogeneidade existente no sumário quantitativo dos resultados dos estudos primários.
 - Análise das fontes ou factores associados à heterogeneidade eventualmente presente (análise de subgrupos, meta-regressão, etc.)
 - Análise de potenciais viéses de publicação.
- Apresentação e adequada discussão das conclusões que é possível retirar a partir dos resultados apresentados.

2. AVALIAÇÃO DA IMPORTÂNCIA CIENTÍFICA E PRÁTICA DOS RESULTADOS

A avaliação da importância dos resultados passa pela adequada interpretação e contextualização das medidas e estatísticas aplicadas na apresentação dos mesmos. Dependendo do tipo de questão de investigação e do tipo de estudo, existe um conjunto de medidas ou estatísticas típicas que é conveniente conhecer e saber interpretar.

Genericamente, a avaliação dos resultados deverá considerar sempre dois aspectos essenciais^[3, 6]:

- A importância prática ou clínica dos resultados, que implica a avaliação da magnitude dos efeitos, medidas ou estatísticas apresentadas, a sua adequada interpretação e a sua devida contextualização clínica.
- A precisão das estimativas dos efeitos, medidas ou estatísticas apresentadas e/ou a avaliação da significância estatística, no caso de serem apresentados resultados relativos a testes de hipóteses.

A necessidade de ser considerado este segundo aspecto está directamente ligada ao facto de, na generalidade dos casos, as estimativas dos parâmetros de interesse obtidas a partir de um determinado estudo serem resultado da análise dos dados de uma amostra, proveniente da população para onde se pretende generalizar os resultados, idealmente seleccionada seguindo um plano de amostragem e tendo um tamanho adequados. Só em situações excepcionais será possível analisar todos os elementos da população que se pretende estudar, portanto, este aspecto é muito relevante. Assim sendo, os erros aleatórios relacionados com o próprio processo de amostragem, e que se manifestam através da incerteza na estimação dos parâmetros de interesse ou através das probabilidades de erros nos testes de hipóteses, estarão sempre presentes e deverão ser considerados através da análise adequada de medidas de precisão das estimativas (ex: intervalos de confiança ou erros padrão das estimativas) e/ou avaliação da significância estatística, no caso de serem

apresentados resultados relativos a testes de hipóteses.

Em seguida são apresentados alguns exemplos de medidas ou estatísticas tipicamente utilizadas na apresentação de resultados, em função do tipo de questão ou estudo considerados. Esta listagem não pretende ser exaustiva, sendo que as medidas e estatísticas para cada tipo de questão ou estudo serão desenvolvidas em maior detalhe nos restantes números desta série, ao serem apresentados exemplos práticos dos mesmos.

Estudos sobre intervenções terapêuticas ou preventivas

Neste tipo de estudos são geralmente apresentadas medidas de efeito que pretendem quantificar o nível de eficácia de uma determinada intervenção no respeitante a uma ou mais variáveis de resultados clínicos de interesse. Alguns exemplos de medidas frequentemente encontradas neste tipo de estudos, e que serão definidas e enquadradas em maior detalhe em números subsequentes desta série, são o risco relativo, a redução absoluta de risco, a redução relativa de risco e o “número necessário para tratar” (*Number needed to treat – NNT*).

Estudos sobre diagnóstico

Os estudos sobre diagnóstico referem-se, frequentemente (mas não exclusivamente), à quantificação da validade de um determinado teste diagnóstico, isto é, comparam os resultados de um determinado teste a avaliar com os resultados de um teste padrão aplicados simultaneamente a uma amostra de indivíduos. Neste contexto, para a apresentação dos resultados neste tipo de estudos são, habitualmente, utilizadas as características operacionais clássicas dos testes diagnósticos (sensibilidade, especificidade e valores preditivos positivos e negativos), assim como um conjunto de outras estatísticas importantes mas um pouco menos conhecidas – *likelihood ratios* positivos e negativos, curvas *ROC*, estatística *C*, *odds ratio* de diagnóstico, etc.

Estudos sobre prognóstico

Os estudos de prognóstico habitualmente estimam a frequência de

determinados fenómenos ao longo da evolução temporal da doença e medem a associação entre o melhor ou pior prognóstico e vários tipos de factores biológicos, psicológicos e sociais. Estes estudos utilizam geralmente medidas de frequência de fenómenos (incidência cumulativa, taxas de incidência, curvas de sobrevivência, probabilidades de sobrevivência, etc.) e medidas de associação típicas (riscos relativos, *odds ratio*, riscos atribuíveis, etc.).

Estudos sobre etiologia ou risco

Os estudos sobre etiologia ou risco geralmente pretendem estimar a associação entre determinados factores e o desenvolvimento de doença e, eventualmente, o impacto desses factores. Neste tipo de estudos são, geralmente, apresentadas medidas de associação, eventualmente ajustadas de forma a corrigir, na medida do possível, os erros sistemáticos relacionados com potenciais variáveis de confusão. Exemplos de medidas de associação e impacto mais frequentemente utilizadas são o risco relativo, o *odds ratio*, o risco atribuível, a percentagem de risco atribuível, a percentagem de risco atribuível populacional, etc.

Estudos de síntese de evidência

Nos estudos de síntese de evidência apresentam-se resultados relativos aos estudos primários fazendo uso das medidas ou estatísticas apropriadas ao tipo de questão ou de estudos que estão a ser considerados. Importa, no entanto, sublinhar que são específicas deste tipo de estudos as medidas de sumário ou agregação de evidência (*pooled estimates*). Estas medidas surgem no contexto da síntese quantitativa dos resultados de um conjunto de estudos primários – meta-análise. Deve ter-se em atenção, no entanto, que esta síntese quantitativa só deverá ser aplicada quando é, de facto, adequado fazê-lo. Neste contexto, o tipo de medidas de sumário apresentadas diz directamente respeito às medidas apresentadas nos estudos primários incluídos.

3. AVALIAÇÃO DA APLICABILIDADE PRÁTICA DA EVIDÊNCIA

A última questão a ser abordada no processo de avaliação crítica da evidência será a questão da aplicabilidade da evidência científica encontrada. Idealmente esta questão só surge após a verificação da adequação metodológica e da importância da evidência científica.

Na verdade, a questão da aplicabilidade não seria obrigatoriamente considerada na avaliação da evidência, no entanto, o objectivo do profissional de saúde será, na maior parte dos casos, a eventual aplicação desta aos seus problemas clínicos e aos seus doentes, logo, a avaliação da aplicabilidade prática é uma questão complementar fundamental neste contexto.

Existem guias e listas de critérios para avaliação da aplicabilidade da evidência^{3,6,17}, tal como para a avaliação da sua qualidade metodológica. A consideração destes critérios pode auxiliar o profissional de saúde a garantir que a utilização que faz da evidência científica, no contexto particular em que está inserido e tendo em conta as características particulares dos seus doentes, é adequada. A título de exemplo apresentam-se em seguida alguns critérios para avaliação da aplicabilidade de evidência científica no âmbito de questões sobre eficácia de intervenções terapêuticas ou preventivas:

- Serão os meus doentes muito diferentes dos participantes do(s) estudo(s) considerado(s)?
- No contexto onde me insiro, estará o tratamento disponível e será aplicável na prática clínica?
- Quais são os potenciais benefícios e malefícios da terapêutica neste doente em particular?
- Quais são as opiniões, valores e expectativas do nosso doente relativamente ao resultado que tentamos prevenir e o tratamento que propomos?

A avaliação da aplicabilidade da evidência científica deverá, idealmente, contemplar três aspectos essenciais – questões biológicas, questões socioeconómicas e questões epidemiológicas¹⁷. O objectivo fundamental desta avaliação será o adequado enquadramento da evi-

dência científica, dadas as características específicas do profissional de saúde, do contexto da sua prática clínica e dos seus doentes.

CONCLUSÃO

Um dos muitos desafios que os profissionais de saúde hoje enfrentam é a necessidade, cada vez mais premente, de integrar a evidência científica de melhor qualidade no apoio à tomada de decisão clínica. A heterogeneidade existente na literatura científica, quanto ao tipo e à qualidade metodológica dos estudos, tornam absolutamente necessária a aplicação de uma abordagem sistemática que permita garantir a qualidade, importância e aplicabilidade da evidência científica relevante e que se pretende aplicar no contexto clínico.

O desenvolvimento de um conjunto de conhecimentos teóricos e competências práticas que permitam realizar uma adequada avaliação crítica da evidência científica é, assim, essencial para qualquer profissional de saúde. Mais importante ainda é a exigência de que esta metodologia seja prática e de fácil implementação de forma a maximizar a probabilidade da sua real utilização.

No próximo número desta série dedicada à pediatria baseada na evidência começar-se-á a apresentação de uma série de exemplos práticos de aplicação dos preceitos e critérios apresentados neste artigo e ao longo dos números anteriores, no contexto dos vários tipos de questões mais frequentemente encontradas na prática clínica diária.

CRITICAL APPRAISAL AND PRACTICAL IMPLEMENTATION OF SCIENTIFIC EVIDENCE

ABSTRACT

Evidence Based Medicine (EBM) is generally defined as the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients. In this section devoted to Evidence Based Paediatrics a series of articles about the conceptual, methodological and operational issues linked to the practice of EBM in the field of

Paediatrics will be published in the following issues of this journal. In the present article theoretical and practical issues related to the critical appraisal and practical implementation of scientific evidence are discussed. The aim of this article is to present a practical and systematic approach for the critical appraisal of evidence, based on three main aspects: (1) assessment of validity and methodological quality of studies; (2) assessment of scientific and practical impact of results; and (3) assessment of applicability of scientific evidence. We present the most common types of research questions and study designs useful in clinical practice; we discuss levels of evidence and grades of recommendation, the concept of systematic error or bias and its classification are discussed; and finally we present a set of assessment criteria for the evaluation of methodological quality, impact and applicability of scientific evidence.

Nascer e Crescer 2008; 17(1): 30-36

BIBLIOGRAFIA

1. Azevedo LF, Costa Pereira A. Pediatria Baseada na Evidência. *Nascer e Crescer Revista do Hospital de Crianças Maria Pia* 2007, 16:29-31.
2. Group E-BMW. Evidence-based medicine. A new approach to teaching the practice of medicine. Evidence-Based Medicine Working Group. *Jama* 1992, 268:2420-2425.
3. Sackett DL, Richardson, W.S., Rosenberg, W., Haynes, R.B. *Evidence Based Medicine*. London, U.K.: Churchill-Livingstone; 1997.
4. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *Bmj* 1996, 312:71-72.
5. Azevedo LF, Costa Pereira A. Pediatria baseada na evidência - Formulação de questões e pesquisa bibliográfica. *Nascer e Crescer Revista do Hospital de Crianças Maria Pia* 2007, 16:135-140.
6. Guyatt G: *Users' guides to the medical literature : essentials of evidence-*

- based clinical practice*. 2nd edn. New York McGraw-Hill Medical; 2008.
7. Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. *Jama* 1995, 274:1800-1804.
 8. Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. *Bmj* 2001, 323:334-336.
 9. Robinson JK, Dellavalle RP, Bigby M, Callen JP. Systematic reviews: grading recommendations and evidence quality. *Arch Dermatol* 2008, 144:97-99.
 10. Schunemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 9. Grading evidence and recommendations. *Health Res Policy Syst* 2006, 4:21.
 11. Sackett DL. Bias in analytic research. *J Chronic Dis* 1979, 32:51-63.
 12. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *Jama* 1993, 270:2598-2601.
 13. Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *Jama* 1994, 271:389-391.
 14. Laupacis A, Wells G, Richardson WS, Tugwell P. Users' guides to the medical literature. V. How to use an article about prognosis. Evidence-Based Medicine Working Group. *Jama* 1994, 272:234-237.
 15. Levine M, Walter S, Lee H, Haines T, Holbrook A, Moyer V. Users' guides to the medical literature. IV. How to use an article about harm. Evidence-Based Medicine Working Group. *Jama* 1994, 271:1615-1619.
 16. Oxman AD, Cook DJ, Guyatt GH. Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group. *Jama* 1994, 272:1367-1371.
 17. Dans AL, Dans LF, Guyatt GH, Richardson S. Users' guides to the medical literature: XIV. How to decide on the applicability of clinical trial results to your patient. Evidence-Based Medicine Working Group. *Jama* 1998, 279:545-549.