

# Como analisar e interpretar um estudo clínico

ÁLVARO AVEZUM

*As recomendações em terapêutica clínica devem ser baseadas em evidências sólidas e robustas da literatura e não em opiniões pessoais baseadas apenas em experiências isoladas.*

*A Cardiologia Baseada em Evidências envolve as seguintes etapas: formular uma questão clara a partir de um problema clínico, pesquisar na literatura artigos clínicos relevantes, avaliar criticamente as evidências quanto à validade e à utilidade clínicas e implementar os achados úteis na prática clínica.*

*A avaliação da literatura no que se refere à terapêutica clínica deve ser realizada de forma crítica e objetiva, utilizando critérios metodológicos para se analisar a validade de resultados e conclusões obtidos. Os seguintes itens devem ser*

*avaliados: validade, dimensão e aplicabilidade dos resultados.*

*Os princípios de epidemiologia clínica devem ser aplicados à prática cardiológica, transformando-nos de espectadores passivos, baseados apenas em opiniões de outros, em médicos praticantes de uma Cardiologia baseada em evidências. A crítica do conhecimento deve ser preferida a seu consumo.*

*Palavras-chave: Cardiologia baseada em evidências, análise crítica da literatura, epidemiologia clínica, estudos clínicos randomizados, terapêutica clínica.*

**HiperAtivo 1998;1:10-9**

*Instituto Dante Pazzanese de Cardiologia*

*Endereço para correspondência:*

*Av. Dr. Dante Pazzanese, 500 — 12º andar — CEP 04012-180 — São Paulo — SP*

## INTRODUÇÃO

A epidemiologia clínica, diferentemente da epidemiologia clássica, refere-se à aplicação dos princípios epidemiológicos no manuseio dos pacientes, detectando e diferenciando, assim, resultados que são clinicamente úteis e verdadeiros daqueles que não o são<sup>(1)</sup>. O conhecimento em Cardiologia pode ser derivado de artigos e livros lidos, de conselhos e sugestões (muitas vezes autoritários) e da experiência pessoal, catalogada atualmente como “na minha experiência”. As fontes dessas informações podem ou não ser verdadeiras, relevantes, válidas e úteis no que diz respeito a uma conduta em Cardiologia. A epidemiologia clínica oferece aos cardiologistas e clínicos critérios específicos e objetivos a serem aplicados na avaliação da validade e utilidade clínica da informação obtida referente à terapêutica clínica.

Virtualmente, a maioria daqueles que praticam Cardiologia (clínica e cirúrgica) obteve seu treinamento na prática e acredita estar praticando a arte da Medicina derivada de concepções, julgamentos e intuições, como também acredita es-

tar praticando a parte científica da Medicina derivada de conhecimento, lógica e experiência prévia. Entretanto, recentemente, temos observado que a ciência é a arte da Cardiologia, permitindo que princípios epidemiológicos (aliados a princípios de bioestatística) sejam aplicados a concepções, julgamentos, intuições, conhecimento, lógica e experiência prévia. Esses princípios epidemiológicos, uma vez aplicados à terapêutica clínica, permitem incrementos substanciais à eficácia e à efetividade das terapêuticas. Como em Medicina, e, particularmente, em Cardiologia, nossa obrigação ética e moral é proporcionar mais benefícios do que riscos aos nossos pacientes, conseqüentemente temos que interpretar crítica e adequadamente as informações que são veiculadas por meio de publicações e apresentações para podermos estar certos de que nossa estratégia terapêutica está correta.

## POR QUE CARDIOLOGIA BASEADA EM EVIDÊNCIAS?

O termo Medicina baseada em evidências (MBE) foi cunhado e primeiramente utilizado na Escola de Medicina da

Universidade McMaster, no Canadá, na década de 1980, para denominar uma estratégia de aprendizado clínico<sup>(2)</sup>. A MBE é um processo que envolve as seguintes etapas: 1) formular questões; 2) localizar as informações disponíveis na literatura; 3) avaliar criticamente as informações relevantes; e 4) utilizar as informações avaliadas criticamente para a decisão clínica<sup>(3)</sup>. Esse processo fundamenta-se nos seguintes itens: 1) as decisões clínicas devem ser baseadas na melhor evidência científica disponível; 2) o problema clínico — ao invés de hábitos e protocolos — deve determinar o tipo de evidência a ser pesquisada; 3) a identificação da melhor evidência significa usar princípios epidemiológicos e bioestatísticos; 4) conclusões derivadas de evidências identificadas e avaliadas criticamente são úteis somente se influenciarem o manuseio de pacientes ou decisões sobre políticas de saúde; e 5) o desempenho de quem pratica MBE deve ser constantemente avaliado. As fontes para obtenção da qualidade das evidências são as seguintes: biblioteca, banco de dados eletrônicos (Medline — 7 milhões de referências), livro-texto (somente se for atualizado regularmente), banco de dados eletrônicos especializados (ACP Journal Club — suplemento bimensal da revista “Annals of Internal Medicine”) e Internet<sup>(4)</sup>. Como inicialmente introduzido em nosso país, gostaríamos de enfatizar novamente o termo Cardiologia baseada em evidências (CBE), que envolve as mesmas etapas descritas anteriormente mas dirigidas para essa área específica. Nos próximos itens desenvolveremos, mais detalhadamente, a racionalidade e a aplicabilidade da CBE.

A prática correta e desejável da Cardiologia envolve a observação das seguintes premissas: efetuar o diagnóstico correto, selecionar o manuseio que proporcione mais benefícios do que riscos e manter-se atualizado no que concerne aos avanços e novidades. Os princípios epidemiológicos e bioestatísticos proporcionam o material necessário para os leitores aumentarem a acurácia e a eficiência de seus esforços clínicos. Torna-se necessário, e urgente, o desenvolvimento de um filtro realmente científico para discernirmos as terapêuticas eficazes daquelas que contêm apenas um perfil cosmético ou mesmo indesejável. Para o desenvolvimento desse filtro, necessitamos incorporar uma nova forma de aprender, pensar e agir. Os princípios da epidemiologia clínica tornam-se, como referido por Sackett e colaboradores<sup>(5)</sup>, “a ciência básica para a Medicina clínica”. O que todas as definições referentes à epidemiologia clínica têm em comum é o respeito pela melhor evidência possível sobre a eficácia e a efetividade dos cuidados médicos e serviços de saúde num período de aumento de complexidade no que se refere ao que nós podemos fazer por nossos pacientes, e também o aumento do reconhecimento de que nós não devemos realizar determinadas condutas. Devido aos espaços conceituais, geralmente substanciais, entre a experiência estruturada da ciência básica e os problemas e discordância muitas vezes intermináveis da assis-

tência clínica aos pacientes, surgiu a necessidade da incorporação da epidemiologia clínica, com o objetivo de estruturar consistente e corretamente a prática clínica. Recentemente, foi iniciado um processo de propagação e disseminação da necessidade de praticarmos uma Medicina baseada em evidências (em nosso caso particular, uma Cardiologia baseada em evidências) em diversos pontos do mundo (Universidade McMaster, Canadá; Universidade Duke, EUA, Universidade de Oxford, Inglaterra, etc.) em detrimento de uma Medicina autoritária, baseada em opiniões de “peritos”, os quais baseiam-se em experiências pessoais (em geral, por definição, experiência pessoal não utiliza grupos controles para analisar a eficácia de uma terapêutica). Apesar de os profissionais que praticam recomendações baseadas em suas experiências, excluindo as evidências da literatura, terem boa intenção, essa é uma prática autoritária, muitas vezes sem concordância entre os vários “peritos” e que podem nem sempre e nem todos estar corretos. A falta de concordância entre as opiniões autoritárias está relacionada não apenas às diferentes condutas diagnósticas e terapêuticas, como também à validade das evidências sobre as quais suas recomendações repousam. O princípio “na minha experiência” foi permitido no passado e não deve mais ser praticado atualmente. Intuitivamente, o raciocínio baseado nesse princípio faz sentido; entretanto, esse raciocínio não passa de uma impressão e deve ser checado por meio de níveis de evidências da literatura. Em muitas situações devemos admitir a incerteza, encorajar o ceticismo e também sermos flexíveis para acomodarmos novas evidências que fatalmente chegam em nossas práticas, as quais podem contradizer o raciocínio e a prática prévios. Evidências que não podem ser reduzidas à ciência de fato (“hard science”) devem ser vistas com incerteza, suspeita e cautela. Em resumo, intuições clínicas, experiências não sistemáticas e racionalidade fisiopatológica não são sinônimos de boa evidência para direcionar condutas. Existem critérios metodológicos para serem seguidos, indistintamente, desde médicos residentes até chefes de departamento, não devendo ser utilizadas interpretações pessoais de resultados de estudos.

Dentre os princípios fundamentais da CBE, podemos citar os seguintes: 1) validade (Podemos confiar na informação?); 2) importância (Caso seja verdadeira, essa informação proporcionará uma diferença importante?); e 3) aplicabilidade (Como podemos usar essa informação?). Em resumo, julgamos que recomendações válidas são oriundas de evidências sólidas e robustas.

## COMO PRATICAR CARDIOLOGIA BASEADA EM EVIDÊNCIAS

A CBE pode ser descrita em quatro etapas: 1) formular uma questão clara a partir de um problema clínico, 2) pesquisar na literatura artigos clínicos relevantes, 3) avaliar criticamen-

te as evidências no que concerne à validade e à utilidade clínica e 4) implementar os achados úteis na prática clínica. A CBE envolve a compreensão da evidência e sua prática está baseada na solidez dessa evidência. A solução prática recomendável é aumentar o conhecimento para se efetuar uma avaliação crítica da literatura para cardiologistas e clínicos que praticam Cardiologia.

Infelizmente, grande proporção de revistas médicas falha em fornecer rigor metodológico relevante e suficiente para responder às diferentes questões criticamente colocadas.

A CBE pode ser aplicada às seguintes áreas: terapêutica, prognóstico, diagnóstico, riscos, revisões sistemáticas (meta-análise), análise econômica, análise de qualidade de vida e tomadas de decisão clínica.

### POR QUE AVALIAÇÃO CRÍTICA DA LITERATURA CARDIOLÓGICA?

Freqüentemente, grande parte dos cardiologistas, ao ler um artigo, o faz de forma superficial, apenas colhendo informações no que diz respeito a resultados e conclusões. Com raras exceções, não é dedicado tempo suficiente ou pelo menos algum tempo à seção de material e métodos, sendo esta considerada complicada, não atraente e, por conseguinte, sem importância. Entretanto, a regra básica e imprescindível é de que não se pode interpretar um resultado ou aceitar e refutar conclusões sem que tenha sido avaliado criticamente o método do artigo. A razão para isso é que um resultado ou conclusão só podem ser considerados válidos se o método for adequadamente considerado. Nesse contexto, uma observação clínica é considerada válida se corresponder ao verdadeiro estado do fenômeno que está sendo avaliado, não podendo ser afetada por erros sistemáticos (viéses) ou efeito do acaso (“play of chance”). Opiniões baseadas em julgamentos de “perito” são em geral consideradas testes fracos e inconsistentes de validade. As conseqüências de se aceitar conclusões espúrias ou conflitantes é de se veicular informações falsas e o pior cenário torna-se a incorporação de terapêuticas clínicas e intervencionistas sem haver confirmação confiável de suas verdadeiras utilidades. Precisamos reunir evidências sólidas para tomarmos decisões e definirmos estratégias terapêuticas quanto ao manuseio de nossos pacientes.

### COMO ANALISAR CRITICAMENTE UM ARTIGO SOBRE TERAPÊUTICA CLÍNICA?

Felizmente, estudos clínicos controlados e randomizados estão sendo realizados com muito mais freqüência e têm demonstrado, de forma clara, a eficácia de alguns tratamentos, como também a ausência de utilidade, ou mesmo o risco, de tantos outros. Entretanto, a avaliação correta de uma terapia requer mais do que randomização. Discutiremos os critérios

que devem ser seguidos para uma avaliação crítica de um artigo, distinguindo, com isso, resultados úteis, conflitantes, confusos, ou mesmo prejudiciais. De forma didática, esses critérios são divididos em três grupos: validade dos resultados, quais foram os resultados e como eles podem ser incorporados ao manuseio de pacientes.

#### Os resultados do estudo são válidos?

Essa questão refere-se à validade ou à acurácia dos resultados e considera se o efeito do tratamento relatado no artigo representa a verdadeira direção e magnitude do efeito do tratamento (Tabela I). Outra maneira de formular e compreender essa questão seria a seguinte: Esses resultados representam uma estimativa sem erros sistemáticos do efeito do tratamento ou apresentam alguma influência de forma sistemática conduzindo a uma falsa conclusão?

#### Os resultados do estudo são válidos?

##### *Critérios primários*

1. A designação dos pacientes para os tratamentos foi realmente randomizada?
2. Todos os pacientes admitidos no estudo foram adequadamente contados e atribuídos à conclusão?
  - O seguimento foi completo?
  - Os pacientes foram analisados nos grupos em que foram originalmente randomizados?

##### *Critérios secundários*

1. Os pacientes, médicos e pessoal envolvido no estudo apresentavam-se sob a condição cega no que diz respeito ao tratamento?
2. Os grupos comparados eram similares no início do estudo?
3. Além da intervenção experimental do estudo, os grupos comparados foram tratados igualmente?

#### Tabela I. Critérios para a avaliação crítica de um artigo sobre terapia. (Adaptado de Guyatt GH e colaboradores<sup>(6)</sup>.)

##### *Critérios primários*

1. A designação dos pacientes para os tratamentos foi realmente randomizada? — O processo de randomização permite estabelecer similaridade entre os dois grupos comparados (por exemplo, tratamento A “versus” controle), no que diz respeito tanto a fatores de risco conhecidos como desconhecidos. Diferentes resultados podem ser obtidos quando o tratamento em comparação é designado de forma randomizada, ou quando essa designação é realizada de forma consciente pelo investigador. Determinados fatores, como severidade da doença, presença de comorbidade e outros fatores prognósticos conhecidos e desconhecidos podem alterar positivamente ou negativamente o efeito de uma terapia em investigação clínica. Como esses

fatores descritos podem influenciar a decisão médica, estudos de eficácia não-randomizados são inevitavelmente limitados quanto a sua habilidade em distinguir se a terapia é útil ou não, como também se existem riscos com a utilização da mesma. A confirmação do exposto acima encontra-se no fato de que a designação de tratamentos por meio de outras maneiras que não a randomização tende a produzir efeitos maiores (e freqüentemente falso-positivos) de tratamento do que em estudos randomizados<sup>(8-11)</sup>. Desde que o tamanho da amostra seja suficiente, a randomização assegura que determinantes prognósticos conhecidos e desconhecidos sejam distribuídos igualmente entre os grupos tratamento e controle. O método de randomização é outro critério a ser avaliado, visto que métodos que utilizam códigos fornecidos por computador ou de maneira automatizada, através de ligação telefônica, onde os critérios de inclusão e exclusão são checados novamente antes de informar qual o tratamento designado, reduzindo as chances de saber qual o tratamento a ser designado antes de se realizar a randomização, são preferíveis àqueles que utilizam envelopes.

2. Todos os pacientes admitidos no estudo foram adequadamente contados e atribuídos à conclusão? — Neste item temos dois componentes:

— O seguimento foi completo? — Todo paciente admitido no estudo deve ser analisado em sua conclusão final. Caso isso não seja feito, ou um número substancial de pacientes seja relatado como “perda de seguimento”, a validade do estudo está aberta para questionamentos. Quanto maior o número de pacientes com perda de seguimento, mais o estudo está sujeito a erros sistemáticos, porque os pacientes que estão perdidos freqüentemente têm prognósticos diferentes daqueles cujo seguimento foi adequadamente realizado. Em casos de perda de seguimento, existem duas possibilidades que podem influenciar os resultados do estudo de forma diferente: a) os pacientes não retornaram para o seguimento porque apresentaram resultados adversos, incluindo óbito; e b) os pacientes estão clinicamente bem e devido a isso não retornaram à visita clínica para serem avaliados. Os leitores podem decidir se a perda de seguimento foi excessiva, assumindo, em estudos com resultados positivos (o tratamento em investigação promove resultados benéficos), que todos os pacientes perdidos no grupo tratamento evoluíram mal (por exemplo, apresentaram óbito) e todos aqueles perdidos no grupo controle evoluíram bem (por exemplo, encontram-se vivos). Recalculando-se os resultados sob essa circunstância proposta, se as conclusões do estudo não forem modificadas, a perda de seguimento não foi excessiva. Entretanto, caso as conclusões se modifiquem, a perda de seguimento deve ser considerada excessiva e o poder de inferência por meio da conclusão obtida torna-se enfra-

quecido e os resultados não são confiáveis<sup>(6)</sup>.

— Os pacientes foram analisados nos grupos em que foram originalmente randomizados? — Na prática clínica, pacientes em estudos randomizados esquecem algumas vezes de utilizar a medicação ou mesmo recusam-se a essa utilização. À primeira vista, os leitores concordarão que pacientes que nunca utilizaram a medicação designada deveriam ser excluídos da análise de eficácia. Freqüentemente, as razões pelas quais os pacientes não utilizam a medicação estão relacionadas ao prognóstico. Alguns estudos clínicos randomizados demonstram que pacientes não-aderentes ao tratamento apresentam pior prognóstico do que aqueles que o são, mesmo considerando-se todos os fatores prognósticos conhecidos e incluindo aqueles pacientes que utilizaram placebo<sup>(12-17)</sup>. A exclusão de pacientes não-aderentes ao tratamento da análise destrói a comparação sem erros sistemáticos proporcionada pela randomização. Esse princípio de atribuir todos os pacientes aos grupos nos quais eles foram originalmente randomizados denomina-se análise de intenção-de-tratar. Essa estratégia preserva o valor da randomização, permitindo a distribuição balanceada de fatores prognósticos nos grupos comparados e, conseqüentemente, o efeito observado será realmente devido ao tratamento designado.

#### *Critérios secundários*

1. Os pacientes, médicos e pessoal envolvido no estudo apresentavam-se sob a condição cega no que diz respeito ao tratamento? — Pacientes que sabem que estão utilizando novo tratamento são mais propensos a ter opinião sobre sua eficácia, como também médicos ou pessoas envolvidas com as avaliações sobre a resposta ao tratamento em questão. Essas opiniões, se otimistas ou pessimistas, podem distorcer sistematicamente os aspectos e o relato dos objetivos do tratamento, reduzindo, conseqüentemente, a confiança nos resultados do estudo. Devido ao fato de pessoas envolvidas na avaliação dos objetivos do estudo não estarem sob a condição cega, podem ser fornecidas diferentes interpretações dos achados durante a realização dos exames em questão, distorcendo, assim, os resultados. A forma mais aconselhável para evitar esse risco à validade é a realização de estudos duplo-cegos (paciente e médico e/ou pessoal envolvido no estudo não são capazes de distinguir o tratamento ativo do placebo).
2. Os grupos comparados eram similares no início do estudo? — Com o objetivo de se assegurarem quanto à validade do estudo, os leitores devem ser informados se os grupos tratamento e controle eram similares com relação aos fatores que podem influenciar e determinar os resultados clínicos de interesse. Esses dados são obtidos pela demonstração das características basais dos grupos tratamento e controle. Embora nunca saibamos se existe similaridade entre os grupos em relação a fatores prognósticos não co-

nhecidos, podemos assegurar-nos de que os fatores prognósticos conhecidos estão adequadamente balanceados. Quando aumentamos o tamanho da amostra, os grupos tornam-se balanceados, tornando mais difícil o efeito do acaso (“play of chance”), sugerindo que um grupo tem aparentemente prognóstico melhor do que o outro. Em estudos adequadamente randomizados, com tamanho de amostra suficiente, sabemos antecipadamente que qualquer diferença entre os grupos tratamento e controle ocorre por acaso, independentemente se houve ou não significância estatística. Entretanto, o ponto em questão é a magnitude dessa diferença. Se for uma grande diferença, a validade do estudo torna-se comprometida. Quanto maior a relação entre fatores prognósticos e resultados e quanto menor o tamanho da amostra do estudo, mais as diferenças entre os grupos enfraquecerão a força de qualquer inferência sobre eficácia. Felizmente, quando os grupos não são similares quanto às características basais, algumas técnicas estatísticas permitem o ajuste dos resultados do estudo para as diferenças entre essas características<sup>(6)</sup>. Obviamente, devemos procurar similaridade entre características relevantes e quando tanto a análise ajustada como a não-ajustada para as características diferentes permitirem a mesma conclusão, admitiremos de forma justificada a validade dos resultados do estudo.

3. Além da intervenção experimental do estudo, os grupos comparados foram tratados igualmente? — As formas de tratamento nos grupos tratamento e controle podem diferir entre si de várias maneiras, e essas diferenças podem também distorcer os resultados do estudo. Intervenções outras além daquela envolvida no estudo são chamadas de co-intervenções, as quais, quando aplicadas diferentemente aos grupos tratamento e controle, podem alterar os resultados. A co-intervenção é um problema sério, principalmente quando a condição duplo-cega está ausente. Como exemplo, podemos ter um estudo avaliando pacientes pós-infarto do miocárdio com determinado tratamento e durante o seguimento um dos grupos (tratamento ou controle) recebe procedimentos de revascularização miocárdica com mais frequência. A co-intervenção, nesse caso, são os procedimentos de revascularização miocárdica, os quais, por terem sido mais utilizados em um grupo, poderão influenciar os resultados favorecendo um dos grupos e, com isto, salientar ou atenuar os efeitos do tratamento em estudo.

Os itens discutidos acima auxiliarão o leitor a determinar se os resultados do artigo são provavelmente válidos ou não. Se os resultados forem considerados válidos, então o leitor poderá prosseguir e considerar a magnitude do efeito do tratamento e sua aplicabilidade.

### Quais foram os resultados?

Caso os resultados sejam válidos e o estudo provavelmente

te conduza a uma avaliação do efeito do tratamento sem erros sistemáticos, então os resultados têm valor para serem examinados. Esta segunda questão considera a dimensão e a precisão do efeito do tratamento (Tabela II). A melhor estimativa desse efeito serão os achados do estudo por si e a precisão da estimativa será superior em estudos com maior número de pacientes.

#### Quais foram os resultados?

1. Qual a dimensão do efeito do tratamento?
2. Qual a precisão da estimativa do efeito do tratamento?
3. As significâncias clínica e estatística foram consideradas?

### Tabela II. Critérios para a avaliação crítica de um artigo sobre terapia. (Adaptado de Guyatt GH e colaboradores<sup>(7)</sup>.)

1. Qual a dimensão do efeito do tratamento? — Existem várias maneiras de se expressar os efeitos de um tratamento:
  - Redução Absoluta de Risco (RAR) — proporção de eventos no grupo controle menos a proporção de eventos no grupo tratamento (por exemplo,  $0,10 - 0,085 = 0,15$ );
  - Risco Relativo (RR) — proporção de eventos no grupo tratamento relativa à proporção de eventos no grupo controle (por exemplo,  $0,085/0,10 = 0,85$ );
  - Redução Relativa de Risco (RRR) — 1 menos proporção de eventos no grupo tratamento relativa à proporção de eventos no grupo controle vezes 100% [ $1 - (0,085/0,10) \times 100\% = 15\%$ ]. Essa é a medida mais comumente utilizada, significando, neste exemplo, que o novo tratamento reduziu o risco de eventos em 15% em comparação com o grupo controle.<sup>(7)</sup>
2. Qual a precisão da estimativa do efeito do tratamento? — A verdadeira redução de risco nunca pode ser conhecida. Tudo o que nós temos é uma estimativa fornecida por estudos clínicos controlados randomizados, e a melhor estimativa do verdadeiro efeito do tratamento é aquela observada nesses tipos de estudos. Temos então um ponto de estimativa devido ao fato da impossibilidade de sermos precisamente corretos. Conseguimos saber os limites (vizinhança) dentro dos quais o verdadeiro efeito provavelmente se encontra. Essa estratégia estatística denomina-se intervalos de confiança (IC)<sup>(18)</sup>. Geralmente, embora arbitrariamente, utilizamos IC 95%, o qual é interpretado como definindo a faixa que inclui a verdadeira RRR 95% das vezes. Significa que a verdadeira RRR encontrar-se-á nesses limites em 95% das vezes. Quanto maior o tamanho da amostra do estudo, maior será o número de eventos e maior será nossa confiança de que a verdadeira RRR (ou outra medida de eficácia) estará próxima daquela que observamos. Torna-se evidente que quanto maior o tamanho da amostra mais estreito é o IC. A precisão de um

resultado de estudo é fornecida pelo IC e não pelo valor de  $p$ , ou seja, mesmo com resultados estatisticamente significantes podemos ter falta de precisão devido a amplos ICs e, conseqüentemente, a importância clínica do resultado estará sujeita a essa falta de precisão. Se por exemplo obtivermos uma RRR de 25% com IC 95% -8% a -40%, significa que se repetirmos o experimento 100 vezes, em 95 das vezes a RRR encontrar-se-á entre 8 e 40. Obviamente, uma RRR de 8% tem importância clínica diferente de uma de 25% ou mesmo de 40%. Quando o artigo não fornece esse importante elemento para avaliarmos a precisão dos resultados obtidos, devemos proceder da seguinte forma: a) caso o valor de  $p$  seja igual a 0,05, significa que o limite inferior do IC para a RRR é zero (não se pode excluir que o tratamento não tenha efeito). Conforme o valor de  $p$  diminui abaixo de 0,05, o limite inferior para a RRR eleva-se acima de zero; b) quando o artigo fornece o erro padrão da RRR (ou do RR), o limite inferior e superior do IC 95% para uma RRR são os pontos estimados mais e menos duas vezes esse erro padrão; c) calcular o IC<sup>(7)</sup>.

### 3. As significâncias clínica e estatística foram consideradas?

— Significância clínica, neste contexto, refere-se à importância da diferença nos resultados clínicos entre os grupos tratamento e controle, sendo geralmente descrita em termos de magnitude de um resultado. Em contraste, significância estatística mostra se as conclusões obtidas pelos autores apresentam probabilidade de ser verdadeiras, independentemente de serem clinicamente importante ou não. Duas questões devem ser formuladas para se avaliar esse item:

— Além de estatisticamente significativa, a diferença encontrada também é clinicamente significativa?

— Se a diferença não é estatisticamente significativa, o estudo apresentava tamanho de amostra suficiente para demonstrar diferença clinicamente importante, caso esta tenha ocorrido?

Como exemplo, imaginamos que um autor conclui que o tratamento A é melhor que B, com um valor de  $p = 0,01$ . Em outras palavras, o valor de  $p$  nos diz o quão freqüente os resultados poderiam ter ocorrido por acaso se a intervenção não fosse diferente do controle. Isso significa que o risco de ter concluído erroneamente que A é melhor que B (quando na verdade não o é) é de apenas 1 em 100. Isso nos parece bom; entretanto, a questão é: Isto é clinicamente importante também? Significância clínica vai além da aritmética e é determinada por julgamento clínico. A medida utilizada para avaliar a significância clínica é o número necessário para tratar (NNT), o qual tem três propriedades para ilustrar sua utilidade clínica:

— enfatiza os esforços utilizados para atingir um alvo tangível de tratamento (auxilia a quantificar e desmistificar a decisão de tratar alguns, mas não todos os pacientes;

— fornece a base para se expressar os custos do tratamento;

— fornece elementos úteis para comparar diferentes tratamentos para diferentes patologias<sup>(5, 7, 19)</sup>.

Após a determinação da magnitude e precisão do efeito do tratamento, a questão final refere-se a como aplicar os resultados do artigo aos pacientes e na prática clínica.

### Como esses resultados podem ser incorporados ao manuseio de pacientes?

Esta questão (Tabela III) apresenta dois pontos importantes:

- Os resultados são aplicáveis aos seus pacientes? Você deveria hesitar em instituir e incorporar o tratamento se seus pacientes não são similares àqueles do estudo, ou se os objetivos onde foi demonstrado benefício não forem suficientemente importantes para seus pacientes.
- Se os resultados são aplicáveis, qual é o impacto do tratamento? O impacto depende dos benefícios e riscos (efeitos adversos e tóxicos) do tratamento e das conseqüências de deixar de tratar os pacientes. Portanto, mesmo um tratamento efetivo poderá não ser utilizado quando o prognóstico do paciente já é bom sem esse tratamento, especialmente quando estiver associado a importantes efeitos adversos e alto custo.

#### Os resultados irão auxiliar-me no manuseio de meus pacientes?

1. Os resultados podem ser aplicados no manuseio de meus pacientes?

(Os pacientes do estudo eram similares aos meus pacientes e o manuseio utilizado pode ser também utilizado na minha prática clínica.)

2. Todos os objetivos clinicamente relevantes foram considerados?

3. Os benefícios do tratamento superam os riscos potenciais e os custos?

### Tabela III. Critérios para a avaliação crítica de um artigo sobre terapia. (Adaptado de Guyatt GH e colaboradores<sup>(7)</sup>.)

- Os resultados podem ser aplicados no manuseio de meus pacientes? — Se os pacientes preencherem os critérios de inclusão e exclusão estabelecidos pelo estudo em questão, não há muita discussão quanto à aplicabilidade dos resultados. Entretanto, se este não for o caso, torna-se necessário outro julgamento e outra análise. Ao invés de argumentarmos que nosso paciente é dois anos mais novo ou o nível de pressão arterial sistólica apresenta diferença de 2 mmHg em relação àquele estipulado pelo estudo, devemos perguntar se existe alguma razão que torna os resultados não aplicáveis aos nossos pacientes. Em geral, uma razão significativa não é encontrada, tornando assim os

- resultados possíveis de serem generalizados para nossos pacientes com razoável confiança. Cuidado especial deve ser tomado quando os resultados globais do estudo não demonstram diferenças significativas e, por meio de análises de subgrupos realizadas posteriormente (sem hipótese pré-especificada), demonstram que determinados subgrupos de pacientes se beneficiam do tratamento. Resumidamente, para generalizarmos os achados de um subgrupo particular aos nossos pacientes, devemos seguir as seguintes etapas em relação ao efeito de tratamento: a) É amplo?; b) É improvável ter sido obtido por acaso?; c) A hipótese foi especificada antes de o estudo ter sido iniciado?; d) Os achados foram também encontrados em outros estudos (replicabilidade)?<sup>(7)</sup>. Os leitores em geral devem ser céticos quanto à aplicação de resultados de análises de subgrupos para seus pacientes.
2. Todos os objetivos clinicamente relevantes foram considerados? — Tratamentos são indicados quando eles fornecem importantes benefícios. Demonstração de que um vasodilatador melhora o débito cardíaco ou de que um antiarrítmico promove supressão de arritmia ventricular não necessariamente fornecem suficiente razão para a prescrição desses tratamentos. O que se torna necessário é a evidência de que esses tratamentos atuem de forma positiva em objetivos que são definitivamente importantes para o paciente, como, por exemplo, redução de hospitalização, redução de mortalidade, ou redução do risco de infarto do miocárdio. A escolha e a utilização de objetivos substitutos (“surrogate endpoints”) devem ser cuidadosas, devendo não apenas refletir o objetivo principal como assegurar que a interferência positiva de um tratamento sobre um objetivo substituto apresente igualmente interferência sobre o que se procura substituir (por exemplo, mortalidade)<sup>(7, 20)</sup>. Um exemplo do risco na escolha de objetivos secundários ocorreu no estudo CAST I e II<sup>(21, 22)</sup>. Os antiarrítmicos (encainida, flecainida e moricizina) utilizados para tratamento de arritmias ventriculares complexas pós-infarto causaram aumento nas taxas de mortalidade, a despeito de se mostrarem eficazes na supressão dessas arritmias. Mesmo quando os autores relatam efeitos favoráveis do tratamento sobre determinados objetivos, os leitores devem certificar-se de que não ocorram efeitos deletérios sobre outros objetivos.
3. Os benefícios do tratamento superam os riscos potenciais e os custos? — Antes da decisão de iniciar determinado tratamento, devem-se considerar os riscos do paciente em apresentar o evento se for deixado sem tratamento. Quanto maior a possibilidade de o paciente apresentar um evento caso não seja tratado, mais provável torna-se o benefício do tratamento para esse paciente e menos pacientes deverão ser tratados para prevenir um evento. O número de eventos adversos decorrentes do uso do tratamento novo

por 1.000 pacientes tratados deve ser contabilizado na decisão final. A incorporação desse risco permite uma análise global e verdadeira da eficácia (por exemplo, a eficácia do ativador tecidual do plasminogênio em comparação com a estreptoquinase deve ser vista como número de vidas salvas mais o número de sangramento intracraniano por 1.000 pacientes tratados). Análises econômicas, principalmente análise de custo-efetividade, a qual considera a eficácia do tratamento e os custos envolvidos na obtenção dessa eficácia, são altamente desejáveis. O princípio de que os recursos serão sempre escassos, utilizado inclusive em países desenvolvidos, deve ser criteriosamente observado em nosso meio, principalmente tendo em vista o panorama econômico de nosso país. Quando se assume que os recursos sempre serão escassos, a utilização eficiente dos tratamentos dependerá da designação desses recursos para determinados tratamentos e situações, os quais devem apresentar evidências sólidas de eficácia terapêutica.

#### Roteiro para recomendações

As recomendações envolvendo a terapêutica em Cardiologia devem seguir os princípios de epidemiologia clínica, os quais permitem a prática do que aqui denominamos Cardiologia baseada em evidências. Procuramos diferenciar e desestimular recomendações fundamentadas em opiniões pessoais, muitas vezes descompromissadas com as verdadeiras e sólidas evidências da literatura. Incentivamos, outrossim, recomendações seguindo princípios de epidemiologia clínica aplicados à Cardiologia definidos objetivamente, os quais encontram-se ao alcance de todos. De forma didática, sugerimos que as recomendações da terapêutica devam se orientar pelas estratégias apresentadas a seguir, com base no desenho do estudo e na proporção de falso-positivos e falso-negativos:

##### a) Tipos de estudos:

- 1 = estudo clínico controlado, randomizado, com baixas taxas de falso-positivos e falso-negativos;
- 2 = estudo clínico controlado, randomizado, com moderadas taxas de falso-positivos e falso-negativos;
- 3 = estudo prospectivo, observacional, utilizando comparações contemporâneas;
- 4 = estudo observacional, utilizando comparações históricas;
- 5 = estudo caso-série.

##### b) Níveis de recomendações:

- A = preferivelmente mais de um estudo, porém pelo menos um estudo tipo 1;
- B = estudos tipo 2;
- C = estudos tipos 3, 4 e 5.

Não deve ser recomendada qualquer terapêutica com base em estudos tipos 3, 4 e 5 (recomendação nível C). Ao invés de gerar recomendações, esses estudos, quando promissores,

	Sim	Não é possível dizer	Não
<b>Os resultados do estudo são válidos?</b>			
A designação dos pacientes para o tratamento foi realmente randomizada?	( )	( )	( )
Todos os pacientes admitidos no estudo foram adequadamente contados e atribuídos à conclusão?	( )	( )	( )
O seguimento foi completo?	( )	( )	( )
Os pacientes foram analisados nos grupos em que foram originalmente randomizados?	( )	( )	( )
Os pacientes, médicos e pessoal envolvido no estudo apresentavam-se sob a condição cega no que diz respeito ao tratamento?	( )	( )	( )
Os grupos comparados eram similares no início do estudo?	( )	( )	( )
Além da intervenção experimental do estudo, os grupos comparados foram tratados igualmente?	( )	( )	( )
<b>Quais foram os resultados?</b>			
Qual a dimensão do efeito do tratamento?	( )	( )	( )
Qual a precisão da estimativa do efeito do tratamento?	( )	( )	( )
As significâncias clínica e estatística foram consideradas?	( )	( )	( )
<b>Os resultados irão auxiliar-me no manuseio de meus pacientes?</b>			
Os resultados podem ser aplicados no manuseio de meus pacientes?	( )	( )	( )
Todos os objetivos clinicamente relevantes foram considerados?	( )	( )	( )
Os benefícios do tratamento superam os riscos potenciais e os custos?	( )	( )	( )

Figura 1. Questionário para avaliação crítica de um artigo sobre terapia em Cardiologia. (Adaptado de Guyatt GH e colaboradores<sup>(6,7)</sup>.)

devem ser encarados e interpretados como geradores de hipóteses a serem testadas adequadamente por meio de estudos clínicos controlados, randomizados, bem conduzidos. Idealmente, as recomendações nível A devem guiar as estratégias terapêuticas em Cardiologia.

## CONSIDERAÇÕES FINAIS

Discutimos, neste artigo, a validade das conclusões, a dimensão, a precisão, os riscos e os custos de efeitos de terapêutica clínica, a generalização e a incorporação na prática clínica dos resultados.

O objetivo básico é o de salientar a necessidade de uma

análise crítica, como também fornecer um roteiro para essa análise da literatura cardiológica. Por meio dessa observação, temos condições de determinar a validade das conclusões de um artigo e, a partir daí, aceitar, incorporar e disseminar as evidências sólidas e válidas encontradas. Quando determinarmos que a qualidade da evidência é insuficiente, devemos observar que qualquer inferência, decisão clínica, ou recomendação estará enfraquecida. Este roteiro torna-se arma fundamental para praticarmos Cardiologia baseada em evidências. Utilizando-se os critérios expostos anteriormente, a análise da literatura cardiológica torna-se mais objetiva, podendo ser praticada por todos, cardiologistas e pessoas que praticam Cardiologia, tornando, conseqüentemente, as impressões sub-

jetivas e as opiniões pessoais não fundamentadas em princípios metodológicos menos freqüentes. Enfatizamos a urgência de praticarmos recomendações fundamentadas em uma literatura válida e confiável, que forneça dados robustos para direcionar nossas condutas. Procedendo assim, evitaremos o erro potencial das recomendações provenientes apenas de experiências e opiniões pessoais, que deixam de lado as evidências muitas vezes já estabelecidas.

Obviamente, a CBE reunirá apoio e atrairá críticas, como

também a compreensão e a incorporação dessa prática poderão não ocorrer em todos os níveis inicialmente, o que é saudável, porque nos ensina a prezar as diferenças existentes. Entretanto, parafraseando o Dr. Curt D. Furberg, durante o último congresso do “American College of Cardiology”, responderemos da mesma maneira: “Em Deus nós confiamos, porém o resto deve mostrar dados” (“In God we trust, the rest must show data” — Curt D. Furberg, Orlando, Março, 1996 — American College of Cardiology Meeting).

## Evidence-based cardiology and critical appraisal of the literature

ÁLVARO AVEZUM

*The evaluation of the cardiology literature should be critically and objectively done by using methodological criteria in order to accept the validity of the results and conclusion. The following items are recommended to be evaluated: validity, dimension and applicability of the results.*

*The recommendation in cardiology therapy should be based upon solid and robust evidence from the literature and not from personal opinion based upon isolated experiences.*

*The evidence-based cardiology can be practiced utilizing the following steps: to formulate a clear question from a clinical problem, to search relevant articles from the literature, to critically evaluate the validity and utility of the evidence and to implement the useful findings in the clinical practice.*

*The clinical epidemiology principles should be applied to the cardiology practice, and then we will practice evidence-based cardiology instead of being passive expectator based on other's opinion.*

*Key words: evidence-based cardiology, critical appraisal literature, clinical epidemiology, randomized clinical trials, cardiology therapy.*

*HiperAtivo 1998;1:10-9*

## REFERÊNCIAS BIBLIOGRÁFICAS

1. Fletcher RH, Fletcher SW, Wagner EH. Clinical Epidemiology: The Essentials. 2<sup>nd</sup> ed. Baltimore: Williams & Wilkins, 1988.
2. Evidence Based Medicine Working Group. Evidence based medicine. JAMA 1992;268:2420-5.
3. Rosemberg W, Donald A. Evidence based medicine: an approach to clinical problem-solving. Br Med J 1995;310: 1122-6.
4. Streiner DL, Norman GR, Blum HM. PDQ Epidemiology. Burlington/Philadelphia: B.C. Decker Inc., 1989.
5. Clinical Epidemiology: a Basic Science for Clinical Medicine. 2<sup>nd</sup> ed. Boston/Toronto/London: Little, Brown and Company, 1991.
6. Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Medicine Working Group. User's Guides to the Medical Literature; II. How to Use an Article About Therapy or Prevention — A. Are the results of the study valid? JAMA 1993;270:2598-601.
7. Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Medicine Working Group. User's Guides to the Medical Literature; II. How to Use an Article About Therapy or Prevention — B. What were the results and will they help me in caring for my patients? JAMA 1994;271:59-63.
8. Sacks HS, Chalmers TC, Smith Jr H. Randomized versus historical assignment in controlled clinical trials. Am J Med 1983;309:1353-61.
9. Chalmers TC, Celano P, Sacks HS et al. Bias in treatment assignment in controlled clinical trials. N Engl J Med 1983;309:1358-61.
10. Colditz GA, Miller JN, Mosteller F. How study design affects outcomes in comparisons of therapy, I: medical. Stat Med 1989;8:441-54.
11. Emerson JD, Burdick E, Hoaglin DC et al. An empirical study of the possible relation of treatment differences to quality scores in controlled randomized clinical trials. Contr Clin Trials 1990;11:339-52.
12. Coronary Drug Project Research Group. Influence of adherence treatment and response of cholesterol on mortal-

- ity in the Coronary Drug Project. *N Engl J Med* 1980;303:1038-41.
13. Asher WL, Harper HW. Effect of human chorionic gonadotropin on weight loss, hunger, and feeling of well-being. *Am J Clin Nutr* 1973;26:211-8.
  14. Hogarty GE, Goldberg SC. Drug and psychotherapy in the aftercare of schizophrenic patients: one-year relapse rates. *Arch Gen Psychiatry* 1973;28:54-64.
  15. Fuller R, Roth H, Long S. Compliance with disulfiram treatment of alcoholism. *J Chronic Dis* 1983;36:161-70.
  16. Pizzo PA, Robichaud KJ, Edwards BK et al. Oral antibiotic prophylaxis in patients with cancer: a double-blind randomized placebo-controlled trial. *J Pediatr* 1983;102:125-33.
  17. Horwitz RI, Viscoli CM, Berkman L et al. Treatment adherence and risk of death after myocardial infarction. *Lancet* 1990;336:542-5.
  18. Altman DG, Gore SM, Gardner MJ et al. Statistical guidelines for contributors to medical journals. In: Gardner MJ, Altman DG, eds. *Statistics with Confidence: Confidence Intervals and Statistical Guidelines*. London, England: British Medical Journal, 1989;pp.83-100.
  19. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *Br Med J* 1995;310:452-4.
  20. Wittes J, Lakatos E, Probstfield J. Surrogate endpoints in clinical trial: cardiovascular diseases. *Statistics in Medicine* 1989;8:415-25.
  21. Echt DS, Liebson PR, Mitchell LB et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo: the Cardiac Arrhythmia Suppression Trial. *N Engl J Med* 1991;324:781-8.
  22. The Cardiac Arrhythmia Suppression Trial II Investigators. Effect of the antiarrhythmic agent moricizine on survival after myocardial infarction. *N Engl J Med* 1992;327:227-33.