

Anna Maria Buehler<sup>1</sup>, Alexandre  
Biasi Cavalcanti<sup>1</sup>, Erica Aranha  
Suzumura<sup>1</sup>, Mariana Teixeira  
Carballo<sup>1</sup>, Otávio Berwanger<sup>1</sup>

## Como avaliar criticamente um ensaio clínico de alocação aleatória em terapia intensiva

*How to assess intensive care randomized trials*

1. Médicos do Instituto de Ensino e  
Pesquisa do Hospital do Coração - IEP-  
HCor - São Paulo (SP), Brasil.

### RESUMO

Ensaio clínico aleatorizado são investigações científicas consideradas padrão-ouro para avaliar intervenções terapêuticas. Ensaio clínico aleatorizado podem examinar e avaliar a segurança e eficácia de novas drogas ou procedimentos terapêuticos ou comparar os efeitos entre duas ou mais drogas ou qualquer outra intervenção. Nesse artigo apresentamos as característi-

cas essenciais e fatores que podem introduzir vieses nesses estudos. Em seguida, apresentamos critérios para avaliação crítica de artigos reportando os resultados de ensaios clínicos aleatorizados e mostramos como interpretar e aplicá-los à prática clínica.

**Descritores:** Avaliação; Ensaio clínico controlado aleatório como assunto; Medicina baseada em evidências/métodos; Unidades de terapia intensiva

---

### INTRODUÇÃO

Os ensaios clínicos aleatorizados são considerados padrão de excelência em estudos científicos que pretendem avaliar o efeito de um tratamento ou qualquer outra intervenção (técnicas ou procedimentos) no curso de uma doença ou situação clínica estabelecida.

É um tipo de estudo muito semelhante aos estudos de coorte prospectivos, com a diferença que o seu delineamento permite eliminar diversos vieses, como, por exemplo, vieses de seleção e fatores de confusão, pois os grupos de tratamento e controle são alocados usando técnicas aleatórias e as características são distribuídas de maneira semelhante em ambos os grupos. Também são estudos submetidos a um controle e gerenciamento mais intensivos.<sup>(1)</sup>

A idéia da distribuição de um tratamento através de aleatorização foi proposta por Fisher, em 1923, aplicada à pesquisa agrícola. A adaptação bem sucedida dos ensaios clínicos aos cuidados com a saúde em humanos se deu somente ao final da década de 40, com o primeiro experimento publicado usando uma tabela de números aleatórios para alocação de sujeitos na pesquisa, realizado pelo Dr. Austin Bradford Hill, da faculdade de Higiene e Medicina Tropical de Londres.<sup>(2)</sup>

### Princípios fundamentais

O quadro 1 resume as principais características dos ensaios clínicos aleatorizados: Trata-se de um estudo experimental, onde se isola a contribuição singular de um fator mantendo-se constante, sempre que possível, os outros determinantes do desfecho. Elege-se o tipo de população alvo da intervenção, estabelecendo-se os critérios de elegibilidade (inclusão e exclusão). Estes critérios podem ser numerosos e

Recebido do Instituto de Ensino e  
Pesquisa do Hospital do Coração -IEP-  
HCor-São Paulo (SP), Brasil.

Submetido em 13 de Janeiro de 2009

Aceito em 23 de Abril de 2009

#### Autor para correspondência:

Otávio Berwanger

Instituto de Ensino e Pesquisa do  
Hospital do Coração (IEP-HCor)  
Rua Abílio Soares, 250 – 12º andar  
CEP: 04005-000 – São Paulo (SP),  
Brasil.

Fone: (11) 3886-4693

FAX: (11) 3886-4695

E-mail: otavioberwanger@yahoo.com.br

### Quadro 1 - Principais características dos ensaios clínicos aleatorizados

Experimentais
Prospectivos
Controlados
Alocação de grupos aleatória e sigilosa
Podem ou não possuir esquema de cegamento

rigorosos quando se quer avaliar a intervenção em uma situação clínica muito específica. Em estudos pragmáticos (*large simple trials*), estes critérios de elegibilidade são bastante resumidos e simples, a fim de aproximar a intervenção avaliada à prática clínica.<sup>(3,4)</sup> Quando muito numerosos e rigorosos, estes critérios podem limitar o recrutamento de pacientes e podem restringir a generalização dos achados para a população em geral. Os critérios são planejados para aumentar a homogeneidade entre os pacientes do estudo, fortalecendo a validade interna.<sup>(5)</sup>

A alocação aleatória permite a geração de grupos verdadeiramente comparáveis, de modo que cada paciente tem a mesma probabilidade de pertencer a um dos grupos (exposto ou não-exposto), desde que todos preencham os critérios de elegibilidade. Todos os fatores relacionados ao prognóstico e ao desfecho tendem a ser igualmente distribuídos nos grupos de comparação.<sup>(1,2)</sup> Desse modo, eventuais diferenças na ocorrência do desfecho entre os grupos experimental e controle podem ser atribuídas à intervenção.

A aleatorização requer cuidados especiais e deve ser gerada por uma técnica adequada. Para a aleatorização ser válida é necessário que cada paciente elegível tenha chance igual de ser alocado para cada um dos grupos do estudo e não pode haver influências dos investigadores. Para tanto, é fundamental que os investigadores não sejam capazes de prever a alocação dos próximos pacientes. É por essa razão que alocar o tratamento conforme o número do prontuário, data de nascimento ou dia da semana não são métodos válidos de randomização.<sup>(6)</sup>

A lista de alocação aleatória geralmente é gerada utilizando softwares apropriados. Outros métodos também considerados válidos são utilização de tabela de números aleatórios, ou mesmo uso de dado ou moeda. Outro item fundamental para garantir a imprevisibilidade da alocação é manter a lista de randomização sob sigilo, ou seja, os investigadores devem primeiro incluir o paciente no estudo e só após o tratamento (experimental ou controle) é definido. Esse critério é denominado de alocação sigilosa e representa o critério metodológico mais importante em um ensaio clínico aleatorizado.<sup>(7)</sup> O método mais efetivo para garantir alocação sigilosa é a randomização central, em que os investigadores cadastram o paciente no estudo via internet ou telefone, e após recebem a alocação do paciente. Outro método aceitável é a utilização de envelopes

lacrados contendo o código de tratamento.

O grupo de tratamento pode ser comparado com um ou mais grupos controles (braços do estudo), que podem utilizar o tratamento corriqueiro da prática clínica, ou placebo (no caso de medicamento), ou até comparar a eficácia e segurança de dois tratamentos distintos para avaliação do desfecho de interesse.

Os ensaios clínicos com alocação aleatória podem referir-se a fármacos, técnicas ou procedimentos e podem ou não possuírem esquema de cegamento.

O cegamento ocorre conjuntamente com a aleatorização e representa o desconhecimento de todos envolvidos na pesquisa (participantes do estudo, investigadores, equipe médica, estatístico) quanto à alocação dos pacientes a um grupo ou a outro. Assim, o estudo não será influenciado por alterações de conduta por parte da equipe médica ou do paciente (efeito Hawthorne).<sup>(8)</sup> O cegamento previne vieses em vários estágios da pesquisa, mas nem sempre ele é possível de ser aplicado. Não se pode ser cego quanto à avaliação de um novo procedimento cirúrgico em unidades de terapia intensiva (UTI), por exemplo.

### AVALIAÇÃO CRÍTICA DOS ENSAIOS CLÍNICOS ALEATORIZADOS

Os intensivistas interessados na leitura de ensaios clínicos aleatorizados devem se atentar para vários aspectos relacionados ao estudo, verificando se o estudo tem validade interna e externa. A validade interna avalia se o estudo conseguiu mensurar adequadamente o que foi proposto, do ponto de vista de relevância clínica e estatística. A validade externa analisa a generalização dos resultados para a prática clínica.

Por ser um tipo especial de estudo de coorte, a maioria dos princípios utilizados para avaliação crítica do estudo são aplicáveis aos ensaios clínicos aleatorizados.<sup>(1)</sup> A seguir, discutiremos os mais relevantes.

#### Validade interna dos resultados

Para uma avaliação crítica da validade interna de um ensaio clínico aleatorizado, vários parâmetros devem ser avaliados, descritos abaixo.

#### Os grupos expostos e não expostos eram semelhantes, exceto pelo fator de exposição?

De uma maneira geral, as possíveis técnicas de aleatorização tendem a evitar que os grupos de pacientes sejam heterogêneos devido à diferenças atribuídas meramente ao acaso.<sup>(9)</sup> Entretanto, quando um estudo possui um tamanho de amostra inadequado, essa homogeneidade não estará garantida. Assim, quanto maior o tamanho da amostra, maior é a garantia que características e fatores relacionados ao des-

fecho do estudo tendam a ser igualmente distribuídos entre os grupos.<sup>(10)</sup> Assim, os intensivistas devem dar preferência a estudos que possuam números de sujeito suficientes, com poder estatístico satisfatório para avaliação do desfecho, em relação à intervenção. Assim, é desejável que uma tabela de comparação das características basais no grupo tratado e no grupo controle seja apresentada para permitir a avaliação de tais características. As características basais referem-se às variáveis demográficas (idade, sexo) e clínicas de interesse, que descrevem os pacientes em estudo. Em um estudo com um tamanho de amostra adequado, estas características são bastante semelhantes.

Havendo diferenças entre os grupos, essas devem ser controladas em análise estatística.

### **O seguimento foi completo?**

Mesmo após a aleatorização, algumas fontes potenciais de perda de segmento devem ser consideradas: alguns pacientes podem não ter a doença como se imaginava no começo do estudo; outros abandonam a pesquisa; não aderem ao tratamento; apresentam reações adversas ou perde-se o contato com eles. Nestes casos, a comparação entre os grupos de estudo torna-se enviesada, mesmo que o delineamento esteja adequado.<sup>(5,11)</sup> Suponha um ensaio clínico onde o desfecho são eventos trombóticos após um ano da alta da internação em UTI. Como profilaxia enquanto internados, o grupo tratamento recebeu uma nova heparina de ultra-baixo peso molecular e o grupo controle recebeu uma outra heparina. Agora suponha que, na determinação do desfecho após um ano, 30% dos pacientes que haviam recebido esta nova heparina não puderam ser localizados. Caso as análises dos resultados demonstrem não inferioridade deste novo medicamento em relação ao grupo controle, este resultado terá que ser analisado com cautela, já que esta perda de seguimento pode ser devido à morte de muitos indivíduos e estes dados não foram contabilizados. Quando os cálculos são refeitos, na verdade o tratamento não demonstra ser mais eficaz que o controle.

O artigo deve disponibilizar as taxas de perda de seguimento para avaliação crítica dos resultados.

### **O estudo foi conduzido segundo o princípio de intenção de tratar?**

É uma metodologia que, na análise primária, inclui nos resultados os dados de todos os pacientes do estudo e que são analisados no braço em que foram inicialmente alocados.<sup>(12)</sup> É um princípio que permite analisar os resultados de acordo com o tratamento designado e não de acordo com o tratamento recebido (estudos explanatórios). Assim, os fatores prognósticos, sabidos ou não, serão, na média, distribuídos igualmente entre os grupos e todos os dados dos pacientes

recrutados serão analisados, independentemente do paciente ter sido seguido até o final do estudo. Também é por este motivo que os resultados de análises secundárias dos estudos devem ser avaliados com atenção, principalmente os que excluem dados de pacientes, já que pode haver perda dos efeitos da aleatorização.<sup>(13)</sup>

### **A intervenção avaliada foi, de fato, a única intervenção relevante no paciente?**

Algumas situações ao longo da condução de um ensaio clínico aleatorizado podem mascarar o real efeito do tratamento investigado. Muitas vezes a utilização de medicações concomitantes ao produto utilizado pode interferir na eficácia do mesmo, quer por questões farmacodinâmicas, quer por questões de efeitos sinérgicos ou antagônicos à medicação utilizada. Tal conceito é definido como co-intervenção. Uma maneira de tentar controlar tal situação é a inclusão da medicação conhecida interferente nos critérios de exclusão ou a anotação periódica de toda medicação concomitante administrada no paciente ao longo do estudo, para serem consideradas no cálculo dos resultados. Caso o paciente esteja utilizando alguma intervenção competitiva previamente à sua entrada no estudo, a influência destas pode ser controlada com a utilização de um período livre de tratamento (*wash-out*) anterior ao estudo.<sup>(14)</sup>

Caso a intervenção investigada já seja conhecida e utilizada na prática clínica, o intensivista deve atentar-se se a mesma não foi utilizada no grupo controle, fora do protocolo de estudo. Esta situação é definida como contaminação em ensaios clínicos aleatorizados, onde se reduz a proporção de grupo controle, uma vez que um percentual deste recebeu a intervenção. A contaminação faz com que os resultados observados tendam a subestimar o verdadeiro efeito do tratamento.<sup>(14)</sup>

### **O estudo foi cego?**

Deve-se dar preferência aos ensaios clínicos aleatorizados que utilizem, quando aplicável, o esquema de cegamento, já discutido previamente. A mensuração do desfecho pode ser influenciada pelo conhecimento da alocação (viés de observação). Tanto os pacientes quanto os investigadores podem ser afetados, em particular para desfechos subjetivos como dor. Adicionalmente, pode haver modificações na condução clínica relacionadas ao conhecimento do tratamento alocado. O cegamento de pacientes, equipe de saúde e investigadores evitam esses vieses.

### **O tamanho da amostra é adequado para a proposta do estudo?**

O ensaio deve recrutar um número suficiente de pacientes para evidenciar a eficácia da intervenção e vários são os

parâmetros utilizados para a determinação deste tamanho de amostra.

A amostra será calculada pensando em minimizar os dois erros  $\alpha$  e  $\beta$ . O erro  $\alpha$  ou tipo I representa o falso positivo, ou seja, a intervenção não é eficaz, mas, pela análise estatística, ela é apontada como eficaz. O erro  $\alpha$  é minimizado pelo nível de significância escolhido, normalmente 5%. Já o erro  $\beta$  ou erro tipo II representa o falso negativo, ou seja, a intervenção é eficaz, mas, pela análise estatística, ela é apontada como não eficaz. O erro tipo  $\beta$  é minimizado indiretamente pelo poder do teste, uma vez que o poder é representado por  $1 - \beta$ . Normalmente deseja-se um poder de, pelo menos, 80%, sendo desejáveis 90%, sempre que possível. O poder do ensaio clínico aleatorizado tem a capacidade de evidenciar a eficácia da intervenção.<sup>(10)</sup>

Assim, para se calcular o tamanho da amostra, deve-se levar em consideração alguns parâmetros: como o desfecho será medido, se quantitativo (numérico) ou se qualitativo (com categorias); nível de significância desejado e poder do teste. Quando o desfecho é quantitativo será necessário ter idéia da magnitude de efeito e desvio padrão. Quando o desfecho é qualitativo é necessário saber qual a proporção de pacientes com desfecho é esperada no grupo da intervenção e qual a proporção de pacientes com desfecho é esperada no grupo controle.<sup>(15)</sup>

### Como apresentar os resultados?

Resultados estatisticamente significantes, ou seja, com probabilidade de significância ( $p$ ) menor que 0,05, podem, muitas vezes, não ser clinicamente relevantes. Para se avaliar a relevância clínica, é necessário considerar a estimativa da magnitude do efeito e qual é a precisão da estimativa.

### Como medir a magnitude de efeito da intervenção?

As medidas de efeito podem ser medidas por uma razão, sendo denominada de medida relativa, ou por uma diferença, considerada medida absoluta. Para medir a magnitude da associação (força do efeito) entre a intervenção e o desfecho (quantas vezes a ocorrência do desfecho é maior no grupo intervenção em relação ao grupo controle) devemos usar medidas de associação relativas. As medidas de associação por diferença avaliam o quanto a frequência de um desfecho no grupo intervenção excede ao grupo controle, ou seja, avaliam qual a incidência do desfecho atribuído a intervenção. As principais formas de medir o efeito são resumidas abaixo, considerando a tabela 2x2 (Tabela 1).<sup>(15)</sup>

- Risco absoluto (RA): É a probabilidade de desenvolver o desfecho em cada grupo. É matematicamente representado por  $a/(a+b)$  e  $c/(c+d)$

- Redução absoluta de risco (RAR): é a diferença de risco

**Tabela 1 – Tabela 2X2**

	Desfecho			Total
		Sim	Não	
Exposição	Sim	a	b	a + b
	Não	c	d	c + d
		a + c	b + d	

Fonte: Traduzida de: Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA. 1995;273(5):408-12.

absoluto entre o grupo controle e entre o grupo tratado. É matematicamente representado por  $RAR = c/(c+d) - a/(a+b)$ . Embora a redução de risco relativo (descrito abaixo) seja o parâmetro mais utilizado para a apresentação dos resultados, a RAR é a medida que tem maior importância clínica, pois é ela quem avalia a eficácia absoluta da intervenção.

- Risco relativo (RR): é o risco de eventos entre pacientes no grupo tratado, relativo ao risco nos pacientes do grupo controle, ou seja,  $RR = [a/(a+b)] / [c/(c+d)]$ . Essa medida nos informa a proporção do risco original que ainda está presente quando os pacientes recebem o tratamento experimental.

- Redução de risco relativo (RRR): É uma estimativa da proporção do risco basal que é removido pelo tratamento experimental. Pode ser calculado de duas maneiras:  $RRR = 1 - RR$  ou  $RRR = RRA/RA$  (no grupo controle). A RRR normalmente tem preferência ao RR na apresentação dos resultados.

- Número necessário para tratar (NNT): é a medida utilizada para avaliar a significância clínica. É matematicamente representada pelo inverso da redução absoluta de risco.  $NNT = 1/RAR$  (um sobre RAR) Expressa o número de pacientes que devem ser tratados por um período de tempo para obter um evento favorável (em caso de tratamento) ou para prevenir um evento desfavorável (em caso de profilaxia). Por exemplo, se uma droga tem um NNT igual a cinco, em relação ao evento morte, significa que cinco pacientes devem ser tratados com ela para que uma morte adicional seja evitada. Considerando um estudo aleatorizado no qual 50% dos participantes morrem no grupo controle e 40% morrem no grupo tratado, a RRR para morte é de 20% e o NNT para evitar a morte será 10 (100/10). Esse tratamento deve ser preferível a outro no qual o NNT para evitar a morte foi de 15. Mas porque diferentes resultados são possíveis, um NNT de 10 não é sempre preferível a um NNT de 15 (se o anterior for angina e o último for qualquer morte, por exemplo). Portanto, um NNT deve sempre vir acompanhado de um resultado claramente indicado e por um período de tempo especificado, para que várias intervenções possam ser comparáveis entre si.

- Número Necessário para Prejudicar (NNP): é calculado

da mesma maneira que o NNT, porém relacionado ao aumento do risco absoluto da intervenção.

Supondo um estudo que queira avaliar uma nova antibioticoterapia na profilaxia de infecção hospitalar, considerem os resultados apresentados na tabela 2 abaixo.

**Tabela 2 – Resultados**

Nova antibioticoterapia	Infecção hospitalar		Total
	Sim	Não	
Sim	5	95	100
Não	15	85	100
	20	180	200

Aplicando as fórmulas descritas anteriormente para estes dados, temos as seguintes interpretações dos resultados:

Risco absoluto (RA): O risco de infecção hospitalar nos pacientes que tomaram a nova terapia foi de 5%. Para os pacientes que tomaram a terapia padrão, este risco foi de 15%.

Redução absoluta de risco (RAR): quem tomou a nova medicação teve reduzido em 10% a probabilidade de ter infecção hospitalar.

Risco relativo (RR): 0,33 é o risco de eventos entre pacientes no grupo intervenção (nova terapia), relativo ao risco nos pacientes do grupo controle. Pode-se dizer que o risco de pacientes do grupo intervenção é igual a 0,33 ou 1/3 do risco dos pacientes no grupo controle.

Redução de risco relativo (RRR): 67% é a redução do risco de ter infecção hospitalar no grupo de pacientes que receberam a nova medicação em relação ao grupo que recebeu o tratamento padrão

Número necessário para tratar (NNT): neste caso é necessário tratar 10 pacientes para se evitar um caso de infecção hospitalar.

### Qual é a precisão do efeito do tratamento?

Devido à variação aleatória, o efeito observado em um estudo provavelmente não será exatamente igual ao do “verdadeiro efeito” da intervenção, entendendo como verdadeiro efeito aquele que observaríamos num estudo similar, mas infinitamente grande. Portanto, é necessária uma medida que nos indique qual o grau de precisão da estimativa do estudo, justamente o que faz o intervalo de confiança.<sup>(7,16)</sup>

O intervalo de confiança de 95% contém o “verdadeiro valor” do efeito com 95% de probabilidade. Quanto mais estreito for o intervalo de confiança (IC) (limites superior e inferior próximos), mais precisão o resultado terá. A probabilidade do valor verdadeiro ficar fora do intervalo é de 5 em 100. Se, por exemplo, obtivermos uma RRR de 25% com IC 95%: 8% a 40% para um determinado evento, significa que se repetirmos o experimento 100 vezes, em 95 das vezes

a RRR encontrar-se-á RRR entre 8 e 40. Obviamente, uma RRR de 8% tem uma importância clínica diferente de uma de 25% ou mesmo de 40%.

Embora o IC seja uma das informações mais importante na avaliação crítica de um ensaio clínico aleatorizado, nem sempre ele é mencionado. Um estudo realizado, de análise de artigos originais com resultados negativos publicados em 1997, nos jornais semanais *British Medical Journal*, *Journal of American Medical Association*, *OK Lancet* e *New England Journal of Medicine* (n=234), teve como objetivo quantificar a proporção de estudos com resultados negativos que comentassem o poder do estudo e o IC do mesmo.<sup>(17)</sup> Este demonstrou que apenas 30% dos estudos comentaram esses parâmetros, sendo que os estudos observacionais os mencionaram menos freqüentemente 15%, 95%CI, (8-21%) que os ensaios clínicos aleatorizados (56%, 95%CI, 46-67%, p< 0,001). Concluíram que os notórios jornais médicos geralmente fornecem informações insuficientes para avaliar a validade dos estudos com resultado negativo.

Todos os artigos deveriam fornecer os IC, entretanto, quando não fornecem, procedemos da seguinte forma: 1) caso o valor de p seja igual a 0,05 provavelmente o limite inferior do IC para RRR será zero (não se pode excluir que o tratamento não tenha efeito). Conforme o valor de p diminui o limite inferior da RRR aumenta. 2) quando o artigo fornece o erro padrão da RRR (ou do RR), os limites inferior e superior do IC 95% para uma RRR são os pontos estimados mais ou menos duas vezes o erro padrão; 3) calcular o IC.

### Validação externa dos resultados

A validade externa de um ensaio clínico aleatorizado está relacionada com a efetividade do estudo, ou seja, a capacidade de generalização dos achados a toda população passível de receber a intervenção estudada. As análises da validade externa envolvem vários aspectos, como variações de paciente, variações etnoculturais, fatores de gravidade, considerações de custo-benefício, risco, infra-estrutura, entre outras. A análise da validade externa só se justifica se a análise da validade interna do estudo for satisfatória. Abaixo, abordaremos as considerações mais importantes.

### Existe a possibilidade de generalizar os resultados do estudo para a população geral e não só para a população-alvo do estudo?

Os critérios de elegibilidade de um ensaio clínico aleatorizado devem indicar a população que o investigador queira inferir.

A situação ideal é a de um estudo livre de vieses (validade interna) e que tenha incluído pacientes comuns da prática clínica (validade externa). Por outro lado, um estudo com

alto risco de viés, mas que inclui uma amostra representativa da população de interesse é pouco útil. Entretanto, muitas vezes nos deparamos com estudos de rigorosos critérios de inclusão/exclusão, que dificultam o recrutamento e limitam a validade externa do mesmo. Nestes casos, deve-se perguntar se, supondo que os resultados sejam verdadeiros, eles podem ser também aplicados aos meus pacientes?<sup>(18,19)</sup>

### Os eventos clinicamente relevantes foram considerados?

O tratamento deve ter impacto sobre os desfechos, que são os eventos clínicos de maior interesse para o paciente e para o médico que o assiste. Os desfechos clínicos são classificados em cinco grupos: desenlace (morte), doença (sintomas, sinais físicos), desconforto (dor, náusea, dispnéia), deficiência funcional (limitação na capacidade de realizar atividades usuais) ou descontentamento (reação emocional à doença ou a seu cuidado).

Relevância clínica vai além da estatística e é determinada por julgamento clínico baseado em evidências estatísticas e por isso é necessário a definição clara de um desfecho relevante. Muitos estudos utilizam desfechos substitutos, que constituem uma medida indireta (marcadores bioquímicos, por exemplo) ou um sinal clínico (redução da despolarização ventricular anormal para redução de arritmias, por exemplo) utilizados em substituição ao desfecho clínico. A vantagem de se utilizar o desfecho substituto é que o tamanho da amostra pode ser minimizado, devido ao fato de tais desfechos geralmente serem mais comuns ou serem variáveis contínuas. Além disso, o desfecho substituto reduz o custo e a duração do estudo. As desvantagens são que, muitas vezes, estas medidas favoráveis à intervenção, que são reportadas primeiramente, escondem os efeitos deletérios do tratamento em outros desfechos, talvez clinicamente mais relevantes, que talvez se manifestassem somente em maior tempo de seguimento.<sup>(3,20)</sup>

A base para o uso do desfecho substituto é que as alterações produzidas pelo tratamento no desfecho substituto devem refletir, de fato, as alterações no desfecho clínico.

### Os benefícios do tratamento irão suplantar os possíveis danos e custos?

Muitas vezes essas questões são resolvidas com análises interinas durante a condução do estudo. A avaliação e acompanhamento de quaisquer reações adversas, muitas vezes, podem gerar a interrupção do estudo antes da data prevista a fim de assegurar segurança ao paciente.<sup>(6,21)</sup>

Ainda, um novo medicamento pode até ter se mostrado eficaz, mas caso exista outro tratamento com uma relação custo-efetividade mais atrativa, não se justifica a utilização desse novo medicamento. O princípio “sempre os recursos são escassos”, utilizado inclusive em países desenvolvidos,

deve ser acentuadamente observado. Quando se assume que os recursos são sempre escassos, a utilização eficiente dos tratamentos dependerá da designação destes recursos para determinados tratamentos e situações, os quais devem apresentar evidências sólidas de eficácia terapêutica.

## CONCLUSÕES

A prática da medicina baseada em evidências já é uma realidade na tomada de decisão de conduta por parte do médico. Hoje em dia não basta apenas sabermos sobre fisiologia e farmacologia e tomarmos como base nossas experiências clínicas e pessoais para termos certeza de que estamos fazendo o melhor pelo paciente.

A análise crítica das evidências exige muito conhecimento por parte do clínico, já que a tomada de decisão baseada em evidências não é trivial. Muito conhecimento é produzido diariamente, mas nem todos os estudos publicados possuem dados ou delineamento de qualidade, para tomarmos como verdadeira tais informações.

Os ensaios clínicos aleatorizados são o padrão-ouro dos estudos para avaliarem intervenções, principalmente tratamento e, quando bem delineados e conduzidos, nos fornecem argumentos fortes e convincentes para o auxílio da tomada de decisão na prática clínica.

Vários são os aspectos que devem ser levados em consideração na leitura de um ensaio clínico aleatorizado e o presente artigo nos permite ter uma noção geral de quão amplo é o conhecimento necessário para a prática da medicina baseada em evidência.

Cabe ao intensivista avaliar criteriosamente a quantidade de estudos conduzidos para uma questão clínica; desta triagem, selecionar os de melhor validade interna e externa e concluir, pela evidência mais convincente e adequada para o perfil de seus pacientes, qual é a melhor conduta a ser seguida.

---

## ABSTRACT

Randomized controlled trials are scientific investigations considered as the gold-standard to evaluate therapeutic interventions. Randomized controlled trials may examine the safety and efficacy of new drugs and therapeutic procedures or compare the effects of two or more drugs or any other intervention. In this article, we present the essential features of these studies, as well as, factors which may bias randomized controlled trials. We also present criteria to critically appraise articles reporting randomized controlled trials, explain how to interpret the results and how to apply them to clinical practice.

**Keywords:** Evaluation; Randomized controlled trials as topic; Evidence-based medicine/methods; Intensive care units

## REFERÊNCIAS

1. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *Am J Ophthalmol.* 2000;130(5):688.
2. Armitage P. The role of randomization in clinical trials. *Stat Med.* 1982; 1(4):345-52.
3. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. *Br J Cancer.* 1976;34(6):585-612.
4. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. *Br J Cancer.* 1977;35(1):1-39.
5. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet.* 2002;359(9308):781-5. Republished in: *Z Arztl Fortbild Qualitatssich.* 2006;100(6):467-73.
6. Schulz KF, Grimes DA. Generation of allocation sequences in randomised trials: chance, not choice. *Lancet.* 2002;359(9305):515-9. Comment in: *Lancet.* 2002;360(9328):258-9; author reply 260. Republished in: *Z Arztl Fortbild Qualitatssich.* 2007;101(6):419-26.
7. Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. *Lancet.* 2002;359(9306):614-8.
8. Schulz KF, Grimes DA. [Epidemiological methods 8: blinded randomized trial: what one covers up is what one obtains]. *Z Arztl Fortbild Qualitatssich.* 2007;101(9):630-7. German.
9. Colditz GA, Miller JN, Mosteller F. How study design affects outcomes in comparisons of therapy. I: Medical. *Stat Med.* 1989;8(4):441-54.
10. Schulz KF, Grimes DA. [Failures in sample size calculation in randomized trial: mandatory and mystical]. *Z Arztl Fortbild Qualitatssich.* 2006;100(2):129-35. Republished from: *Lancet.* 2005;365(9467):1348-53. German.
11. Sackett DL. Participants in research. *BMJ.* 2005;330(7501):1164. Comment in: *BMJ.* 2005;331(7508):109-10. Comment on: *BMJ.* 2005;330(7501):1175.
12. Lewis JA, Machin D. Intention to treat--who should use ITT? *Br J Cancer.* 1993;68(4):647-50.
13. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet.* 2002;359(9308):781-5. Republished in: *Z Arztl Fortbild Qualitatssich.* 2006;100(6):467-73.
14. Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet.* 2002;359(9302):248-52. Comment in: *Lancet.* 2002;360(9328):258.
15. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA.* 1995;273(5):408-12. Comment in: *JAMA.* 2001;286(20):2546-7.
16. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med.* 1988;318(26):1728-33.
17. Hebert RS, Wright SM, Dittus RS, Elasy TA. Prominent medical journals often provide insufficient information to assess the validity of studies with negative results. *J Negat Results Biomed.* 2002;1:1.
18. Sackett DL. Evidence-based medicine and treatment choices. *Lancet.* 1997;349(9051):570; author reply 572-3. Comment on: *Lancet.* 1997;349(9045):126-8.
19. Sackett DL. Evidence-based medicine. *Spine.* 1998;23(10):1085-6.
20. Miller JN, Colditz GA, Mosteller F. How study design affects outcomes in comparisons of therapy. II: Surgical. *Stat Med.* 1989;8(4):455-66.
21. Montori VM, Devereaux PJ, Adhikari NK, Burns KE, Egger CH, Briel M, et al. Randomized trials stopped early for benefit: a systematic review. *JAMA.* 2005;294(17):2203-9. Comment in: *Crit Care.* 2007;11(1):305. *JAMA.* 2005;294(17):2228-30.