

Medicina baseada em evidências e hipertensão arterial

Medicine based on evidences and hypertension

Coordenador: *Álvaro Avezum*

Os dez mandamentos do ensaio clínico randomizado – princípios para avaliação crítica da literatura médica

The ten commandments of the randomized clinical trial – principles for critical appraisal of the medical literature

Otávio Berwanger¹, Hélio Penna Guimarães¹, Álvaro Avezum², Leopoldo Soares Piegas³

INTRODUÇÃO

Ensaio clínico randomizado representa, dentre os estudos individuais, o padrão-ouro para a avaliação de intervenções terapêutico-preventivas. Essa estratégia de pesquisa, proposta pela primeira vez em 1948, por Hill, encontra-se entre as mais importantes descobertas da medicina no século XX¹⁻³. Em um ensaio clínico randomizado, a designação de pacientes para os tratamentos ou exposições encontra-se sob o controle do investigador e por isso esse é considerado um delineamento do tipo experimental. Nesse tipo de estudo, os pacientes são designados de maneira randomizada ou aleatória para qualquer uma das intervenções sob estudo. A randomização procura assegurar que os grupos fiquem balanceados, tanto para características conhecidas quanto desconhecidas, de forma que a única diferença entre os grupos sejam as intervenções experimental e controle^{4,5}. Assim, a randomização representa o grande diferencial sobre os demais delineamentos (estudos observacionais e quase-experimentos) por permitir, quando adequadamente implementada, um método extremamente eficaz para o controle de erros sistemáticos⁶⁻⁸. Após a randomização, os grupos são seguidos por período de tempo específico e analisados em termos de desfechos de interesse definidos pelo protocolo do estudo (exemplo: óbito, reinfarto, hospitalização por insuficiência cardíaca etc.). Dado que a única diferença entre os grupos são os tratamentos experimentais e controle, caso ocorrerem diferenças no evento desses desfechos ao final do estudo, essas diferenças podem ser atribuídas ao tratamento experimental.

Visto que a hipertensão arterial representa uma das doenças mais bem estudadas, por meio de ensaios clínicos randomizados, na seqüência, apresentamos 10 mandamentos ou princípios fundamentais que podem ser utilizados pelo médico a fim de julgar se a evidência disponível possui qualidade satisfatória para ser incorporada na prática clínica diária.

MANDAMENTO I

A MANUTENÇÃO DO SIGILO DA LISTA DE RANDOMIZAÇÃO REPRESENTA O ASPECTO METODOLÓGICO MAIS IMPORTANTE DE UM ENSAIO CLÍNICO RANDOMIZADO

Não basta grande número de pacientes para obtermos uma randomização adequada. Para tanto, se faz necessário a manutenção do sigilo da lista de alocação, que representa a principal característica de um ensaio clínico randomizado⁹⁻¹¹. Dentro da pesquisa clínica, o conceito de randomização envolve necessariamente “incapacidade de prever”, ou seja, a pessoa responsável por alocar os sujeitos de pesquisa aos diferentes tratamentos, não pode, sob hipótese alguma, prever para qual tratamento o paciente será alocado. Dessa forma, sem a adequada manutenção do sigilo da lista de alocação, os benefícios da randomização são perdidos. Dessa forma, após ser gerada por um programa para computador ou por meio de uma tabela de números aleatórios, a lista de randomização deve ser mantida em sigilo. A fim de respeitar esse critério, é necessário lançar mão de métodos que utilizam códigos fornecidos por computador ou de maneira automatizada (através de ligação para

Recebido: 05/12/2005 Aceito: 30/01/2006

1 Médico, Coordenador e Investigador de Estudos Clínicos da Divisão de Pesquisa, Instituto Dante Pazzanese de Cardiologia (São Paulo-SP). Doutor em Epidemiologia pela Faculdade de Medicina da Universidade Federal do Rio Grande do Sul.

2 Médico, Coordenador e Investigador de Estudos Clínicos – Divisão de Pesquisa, Instituto Dante Pazzanese de Cardiologia (São Paulo-SP). Médico Assistente da Disciplina de Clínica Médica da Universidade Estadual de São Paulo.

3 Diretor da Divisão de Pesquisa do Instituto Dante Pazzanese de Cardiologia (São Paulo-SP). Doutor em Cardiologia pela Faculdade de Medicina da Universidade de São Paulo.

4 Diretor do Instituto Dante Pazzanese de Cardiologia (São Paulo-SP). Professor Livre-Docente em Cardiologia pela Faculdade de Medicina da Universidade de São Paulo.

Correspondência para: Dr Otávio Berwanger e Dr Hélio Penna Guimarães. Divisão de Pesquisa – Instituto Dante Pazzanese de Cardiologia. Av. Dr Dante Pazzanese 500, 13º andar – 04012-909 – São Paulo – SP. E-mail: otavioberwanger@yahoo.com.br

central telefônica ou por meio de *web-based randomization*), onde os critérios de inclusão e exclusão são checados novamente antes de informar qual o tratamento designado. Assim sendo, há redução das chances de saber qual o tratamento a ser designado antes de se realizar a randomização. Por outro lado, métodos que utilizam envelopes não-selados, não-seriados ou não-opacos, divulgação da lista de randomização em estudos abertos, bem como alocação baseada em sorteio do tipo “par ou ímpar” (utilizando, por exemplo, a data de nascimento ou último dígito do número do prontuário do paciente) não são considerados métodos válidos de randomização, justamente por permitirem que o sigilo da lista de alocação seja quebrado. Um clássico exemplo de quebra do sigilo da lista de randomização ocorreu no estudo CAPPP, no qual 10.985 pacientes foram randomizados para receber captopril 50 mg, uma a duas vezes por dia, ou drogas convencionais (beta-bloqueadores ou diuréticos)¹². Ao final de cinco anos, não foram encontradas diferenças significativas em relação a eventos cardiovasculares maiores, contudo a incidência de AVC foi superior nos pacientes que receberam captopril. Neste estudo, o uso de envelopes não-opacos, não-lacrados e não-seriados permitiu quebra do sigilo da lista de randomização, de forma que ocorreu alocação sistemática dos indivíduos aos tratamentos. Como resultado, a perda da integridade da randomização gerou desbalanço entre as características basais dos pacientes, sendo que o grupo do captopril apresentou níveis pressóricos mais elevados e maior número de diabéticos. Este grave erro metodológico compromete a interpretação dos resultados do estudo.

MANDAMENTO II

SEMPRE QUE POSSÍVEL, PACIENTES, MÉDICOS E INVESTIGADORES DEVEM DESCONHECER A IDENTIDADE DOS TRATAMENTOS

Pacientes que têm conhecimento sobre a identidade do tratamento que estão utilizando são mais prováveis de terem uma 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 estudo¹³⁻¹⁵. 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 à realização de estudos cegos (idealmente, paciente e médico e/ou pessoal envolvido no estudo não são capazes de distinguir o tratamento ativo do placebo). Atualmente, o termo duplo-cego está em desuso, uma

vez que seu conceito não exige uniformidade na interpretação de seu significado entre diferentes pesquisadores¹⁶. Dessa forma, o *Consort Statement* recomenda utilizar apenas a terminologia “estudo cego”, descrevendo quais das partes envolvidas desconhecem a natureza das intervenções em estudo¹⁷.

MANDAMENTO III

O SEGUIMENTO DEVE SER COMPLETO

Todo paciente que foi admitido no estudo deve ser analisado na 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 àqueles nos quais o seguimento foi adequadamente realizado. Em casos de perda de seguimento, duas possibilidades, influenciando os resultados do estudo de forma diferente, podem ocorrer: 1) os pacientes não retornaram para o seguimento porque apresentaram resultados adversos, incluindo óbito; 2) os pacientes estão clinicamente bem e devido a isso não retornaram à visita clínica para serem avaliados. 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 (exemplo: apresentaram óbito) e todos aqueles perdidos no grupo controle evoluíram bem (exemplo: encontram-se vivos). Recalculando 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, através da conclusão obtida, torna-se enfraquecido e os resultados não são confiáveis.

MANDAMENTO IV

A ANÁLISE DEVE SER POR INTENÇÃO-DE-TRATAR

Na prática clínica, pacientes em estudos randomizados esquecem algumas vezes de utilizar a medicação ou mesmo recusam-se à essa utilização. À primeira vista, pareceria lógico 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 todos os fatores prognósticos conhecidos e incluindo aqueles pacientes que utilizaram placebo.

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-tratar¹⁸. Essa estratégia preserva o benefício da randomização, permitindo a distribuição balanceada de fatores prognósticos nos grupos comparados e, conseqüentemente, o efeito observado será realmente devido ao tratamento designado.

MANDAMENTO V

GRUPOS DEVEM SER TRATADOS IGUALMENTE ALÉM DA INTERVENÇÃO EXPERIMENTAL DO ESTUDO

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 cointervenções e quando aplicadas diferentemente aos grupos tratamento e controle podem alterar os resultados. A cointervenção 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) receber procedimentos de revascularização miocárdica com mais frequência. A cointervenção 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 isso salientando ou atenuando os efeitos do tratamento em estudo.

MANDAMENTO VI

A QUALIDADE METODOLÓGICA INFLUENCIA DIRETAMENTE O RESULTADO DE UM ENSAIO CLÍNICO RANDOMIZADO (INDEPENDENTE DE QUALQUER EFEITO METODOLÓGICO DAS INTERVENÇÕES EM ESTUDO)

O não respeito a critérios metodológicos (quebra do sigilo da lista de randomização, ausência de cegamento, perdas no seguimento e ausência de análise por intenção de tratar e grupos de forma não-uniforme) podem alterar de forma significativa os resultados de um ensaio clínico randomizado. Nesse sentido, estão disponíveis evidências empíricas na literatura demonstrando que, independente de qualquer efeito biológico das intervenções em estudo, a presença de qualquer dessas limitações metodológicas, na maior parte das vezes, tende a levar a uma superestimativa do efeito de uma intervenção, contudo a direção dos efeitos é imprevisível, podendo, em algumas ocasiões, ocorrer exatamente o oposto, ou seja, subestimativa de efeito^{19;20} (Tabela 1).

Tabela 1. Influência da qualidade metodológica no resultado de ensaios clínicos randomizados.

Limitação metodológica (Inadequado versus adequado)	Efeito nos resultados
Método de randomização inadequado	Superestimativa de 20%
Quebra do sigilo da lista de alocação	Superestimativa de 30%
Ausência de cegamento de pacientes e médicos	Superestimativa de 18%

MANDAMENTO VII

A DECISÃO CLÍNICA DEVE SER BASEADA EM DESFECHOS CLINICAMENTE RELEVANTES E NÃO EM DESFECHOS SUBSTITUTOS

Tratamentos são indicados quando fornecem importantes benefícios para o paciente, para o médico e para a sociedade, como por exemplo, redução de mortalidade total e da mortalidade cardiovascular, redução do risco de infarto do miocárdio e redução do risco de AVC. Esses são os chamados desfechos clinicamente relevantes ou, simplesmente, eventos cardiovasculares maiores. Dessa forma, o principal objetivo do tratamento do paciente com hipertensão arterial é justamente a redução desses eventos. A redução dos níveis de pressão arterial sistólica e diastólica, a melhora da fração de ejeção, a melhora de parâmetros laboratoriais (glicemia, perfil lipídico) devem ser considerados desfechos substitutos (*surrogate endpoints*). Os desfechos substitutos são importantes, sem dúvida, mas servem apenas para gerar novas hipóteses, nunca para mudar a prática clínica.

MANDAMENTO VII

GRANDE NÚMERO DE PACIENTES E (PRINCIPALMENTE) DE EVENTOS É NECESSÁRIO PARA DEMONSTRAR REDUÇÕES PLAUSÍVEIS (MODERADAS) DE EVENTOS CARDIOVASCULARES

A maioria das intervenções terapêuticas em cardiologia promove reduções moderadas e plausíveis de mortalidade, ou seja, a verdadeira redução de risco encontra-se provavelmente entre 10% e 25%, em vez de 40% a 60% – princípio da moderação. Assim, os resultados dos estudos clínicos disponíveis, envolvendo o número adequado de pacientes demonstram que, em sua maioria, tratamentos promovem reduções moderadas, porém, plausíveis de eventos cardiovasculares maiores e de mortalidade^{21;22}. Devido à alta prevalência e incidência da doença cardiovascular, essas reduções de magnitude moderada podem evitar ou prorrogar dezenas de milhares de eventos anualmente, o que é altamente desejável do ponto de vista de saúde coletiva. Na tabela 2, adaptada da série de revisões narrativas publicadas por Califf e DeMets, é possível notar que a maioria das intervenções em cardiologia estão de acordo com os princípios da moderação²³⁻²⁶.

A fim de detectar essas reduções moderadas com poder estatístico adequado, é necessário o tamanho de amostra não inferior a 6.000 pacientes e/ou pelo menos a observação de no mínimo 350 a 650 eventos (Tabela 3)²⁷⁻²⁹.

MANDAMENTO IX

A SIGNIFICÂNCIA CLÍNICA E ESTATÍSTICA DEVEM SER CONSIDERADAS
Relevância clínica nesse contexto refere-se à importância da diferença nos resultados clínicos entre os grupos tratamento e controle, sendo geralmente descrita em termos de magnitude de resultado. Em contraste, significância estatística mostra se as conclusões obtidas pelos autores apresentam probabilidade de serem verdadeiras, independentemente de serem clinicamente importantes ou não. Duas questões devem ser formuladas para se avaliar esse item: 1) além de estatisticamente significativa, a diferença encontrada também é clinicamente significativa?; 2) se a diferença não é estatisticamente significativa, o estudo apresentava tamanho de amostra suficiente para demonstrar

diferença clinicamente importante, caso essa tenha ocorrido? Como exemplo, imaginamos que um autor conclui que o tratamento A é melhor do que B, com um valor de $p = 0,01$. Em outras palavras, o valor de p diz 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 do que B (quando na verdade não o é) é de apenas 1 em 100. Isso parece atrativo, entretanto a questão fundamental é: "isso 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 têm três propriedades para ilustrar sua utilidade clínica: a) enfatiza os esforços utilizados para atingir alvo tangível de tratamento (auxilia a quantificar e desmistificar a decisão de tratar alguns, mas não todos os pacientes); b) fornece a base para se expressar os custos do tratamento; c) fornece elementos úteis para comparar diferentes tratamentos para diferentes doenças. O NNT expressa quantos pacientes necessitam ser tratados por dado período de tempo para se evitar certo desfecho³⁰. Quanto menor o NNT, mais importante é o efeito do tratamento. O NNT pode ser facilmente obtido dividindo-se 100 pela RAR (em porcentagem). Por exemplo, se em um ensaio clínico randomizado é obtida uma RAR de 2%, o NNT pode ser facilmente obtido pelo cálculo $100/2 = 50$. (Interpretação: "é necessário tratar 50 pacientes com a intervenção experimental, por cinco anos para que um paciente adicional deixe de sofrer o desfecho em estudo"). É bem definido o conceito de que quanto maior o risco de eventos, maior o potencial benefício do tratamento. No caso da hipertensão arterial sistêmica, amplo conjunto de evidências demonstrou que o risco de eventos cardiovasculares é contínuo e aumenta de forma exponencial com o aumento da pressão arterial. Logo, é esperado que o benefício do tratamento anti-hipertensivo seja maior no indivíduo com níveis pressóricos mais elevados. Nesse sentido, o NNT é medida melhor do que RRR para realizar decisões sobre o manejo dos pacientes, uma vez que a RRR não reflete a magnitude do risco absoluto. Um artigo publicado por Cook e Sackett ilustra esse conceito através de uma revisão sistemática que testou o efeito do tratamento anti-hipertensivo em pacientes com níveis de pressão acima e abaixo de 110 mmHg, respectivamente³⁰. A partir desses dados, nota-se que os resultados expressos sob a forma de RRR sugerem falsamente que o benefício do tratamento é semelhante nos dois grupos (RRR = 40%). Por outro lado, quando os resultados são expressos sob a forma de NNT, fica claro que o benefício é maior (menor NNT) no grupo com níveis pressóricos mais elevados (NNT = 13) do que no grupo com pressão diastólica menor ou igual a 110 mmHg (NNT = 167).

Tabela 2. Magnitude da redução de risco relativo com o emprego de intervenções cardiovasculares.

Intervenção	Cenário	Nº de pacientes	RRR
Aspirina	IAM	18.773	23%
Trombolíticos	IAM	58.000	18%
Beta-bloqueadores	IAM	29.970	13%
IECAIAM	101.000	6,5%	
Aspirina	Prevenção secundária	54.360	15%
Estatinas	Prevenção secundária	20.312	23%
IECAICC	7.105	23%	
Beta-bloqueadores	ICC	12.385	26%
Espironolactona	ICC	1.663	30%

Tabela 3. Número de pacientes de eventos para obtenção de poder estatístico adequado (considerando a taxa de eventos no grupo controle de 10% e a RRR de 25%).

Número de eventos	Número de pacientes	Chance de erro tipo II	Comentários sobre tamanho da amostra
0 - 50	< 500	> 0,9	Claramente inadequada
50 - 150	1.000	0,7 - 0,9	Provavelmente inadequada
150 - 300	3.000	0,3 - 0,7	Possivelmente inadequada
350 - 650	6.000	0,1 - 0,3	Provavelmente adequada
> 650	10.000	< 0,1	Adequada

MANDAMENTO X

A ESTIMATIVA DE EFEITO GLOBAL SEMPRE É MAIS PRECISA E CONFIÁVEL DO QUE AS ESTIMATIVAS DE EFEITO OBTIDAS EM SUBGRUPOS

Além do resultado global, artigos sobre ensaios clínicos randomizados, usualmente, também apresentam as chamadas análises de subgrupos, isto é, o resultado em pacientes com diferentes características a exemplo de gênero, gravidade da doença, presença de comorbidades, uso concomitante de outras medicações, dentre outros. Para ser válida, uma análise de subgrupo deve ser definida *a priori*, deve ser restrita a poucos subgrupos, esses com número de pacientes e eventos relativamente adequado sempre que possível e ser interpretada à luz da totalidade das evidências prévias, a fim de estabelecer se os achados são consistentes e plausíveis do ponto de vista clínico-epidemiológico. A análise de subgrupo serve, no máximo, como geração de hipótese para novos estudos, nunca devendo ser interpretada como achado definitivo. Dessa forma, a estimativa global do estudo, por incluir um maior número de pacientes e, principalmente, de eventos, sempre é mais precisa e confiável do que as estimativas de efeito obtidas nos diferentes subgrupos, devendo ser aquela valorizada primariamente. Um exemplo clássico é a análise de subgrupos do ISIS-2, apresentada de forma proposital por Richard Peto para demonstrar que as análises de subgrupo são “verdadeiras máquinas de gerar resultados falso-positivos”. O estudo ISIS-2 demonstrou de forma robusta e satisfatória que aspirina e trombolíticos reduzem mortalidade em IAM³¹. Quando indagados sob a ausência de análises de subgrupo no relato do estudo, os autores argumentaram que a estimativa global era a de que realmente importava do ponto de vista de decisão clínica. Após a insistência pela apresentação da análise de subgrupos por parte de terceiros, os autores fizeram questão de demonstrar o risco de associações espúrias ao relatar que o efeito da aspirina, na redução da mortalidade, poderia variar de acordo com o signo astrológico (Tabela 4). Assim, no subgrupo de pacientes dos signos de gêmeos ou libra, a aspirina não possuiria benefício, enquanto que em pacientes dos outros signos o uso dessa intervenção seria benéfico!

Tabela 4. Análise de subgrupos do estudo ISIS-2 de acordo com o signo astrológico.

Signo	Taxa de eventos – aspirina	Taxa de eventos – Placebo	Valor de P
Libra	150 (11,1%)	147 (10,2%)	0,5
Todos os demais	654 (9%)	869 (12,1%)	< 0,0001

CONSIDERAÇÕES FINAIS

Na busca de uma prática baseada em evidências, é necessário que médicos possuam capacidade de avaliar criticamente a evidência disponível. Nesse sentido, ao avaliar um artigo sobre tratamento (particularmente um ensaio clínico randomizado) na área de hipertensão arterial, o médico deve julgar aspectos de qualidade metodológica (manutenção do sigilo da lista de randomização, cegamento, minimização das perdas e análise por intenção de tratar), poder estatístico (dado, principalmente, pelo número de eventos), efeito da intervenção sobre desfechos clinicamente relevantes, principalmente levando em conta a significância clínica (por meio de NNT) além da estatística. Adicionalmente, a análise de subgrupos deve ser valorizada do ponto de vista de geração de novas hipóteses, sendo a estimativa global do estudo aquela que deve direcionar a tomada de decisão clínica.

REFERÊNCIAS

- Hill AB. Suspended judgment. Memories of the British Streptomycin Trial in Tuberculosis. The first randomized clinical trial. *Control Clin.Trials* 1990;11:77-9.
- Hill AB. The clinical trial. *N.Engl.J.Med.* 1952;247:113-9.
- Yusuf S, Collins R, Peto R. Why do we need some large, simple randomized trials? *Stat.Med.* 1984;3:409-22.
- Altman DG, Bland JM. Statistics notes. Treatment allocation in controlled trials: why randomise? *BMJ* 1999;318:1209.
- Schulz KF, Grimes DA. Generation of allocation sequences in randomised trials: chance, not choice. *Lancet* 2002;359:515-9.
- Peto R, Pike MC, Armitage P *et al.* Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. *Br.J.Cancer* 1976;34:585-612.
- Peto R, Pike MC, Armitage P *et al.* Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. *Br.J.Cancer* 1977;35:1-39.
- Peto R. Clinical trial methodology. *Biomedicine.* 1978;28 Spec No:24-36.
- Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. *Lancet* 2002;359:614-8.
- Schulz KF. Assessing allocation concealment and blinding in randomized controlled trials: why bother? *ACP J.Club.* 2000;132:A11-A12.
- Altman DG, Schulz KF. Statistics notes: Concealing treatment allocation in randomised trials. *BMJ* 2001;323:446-7.
- Hansson L, Lindholm LH, Niskanen L *et al.* Effect of angiotensin-converting-enzyme inhibition compared with conventional therapy on cardiovascular morbidity and mortality in hypertension: the Captopril Prevention Project (CAPPP) randomised trial. *Lancet* 1999;353:611-6.
- Day SJ, Altman DG. Statistics notes: blinding in clinical trials and other studies. *BMJ* 2000;321:504.
- Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ *et al.* Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin.Trials* 1996;17:1-12.
- Schulz KF, Chalmers I, Altman DG. The landscape and lexicon of blinding in randomized trials. *Ann.Intern.Med.* 2002;136:254-9.
- Devereaux PJ, Manns BJ, Ghali WA *et al.* Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. *JAMA* 2001;285:2000-3.
- Altman DG, Schulz KF, Moher D. Turning a blind eye: testing the success of blinding and the CONSORT statement. *BMJ* 2004;328:1135.
- Montori VM, Guyatt GH. Intention-to-treat principle. *CMAJ.* 2001;165:1339-41.
- Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *BMJ* 2001;323:42-6.
- 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;273:408-12.

21. Collins R, Gray R, Godwin J, Peto R. Avoidance of large biases and large random errors in the assessment of moderate treatment effects: the need for systematic overviews. *Stat.Med.* 1987;6:245-54.
22. Collins R, MacMahon S. Reliable assessment of the effects of treatment on mortality and major morbidity, I: clinical trials. *Lancet* 2001;357:373-80.
23. DeMets DL, Califf RM. Lessons learned from recent cardiovascular clinical trials: Part I. *Circulation* 2002;106:746-51.
24. DeMets DL, Califf RM. Lessons learned from recent cardiovascular clinical trials: Part II. *Circulation* 2002;106:880-6.
25. Califf RM, DeMets DL. Principles from clinical trials relevant to clinical practice: Part I. *Circulation* 2002;106:1015-21.
26. Califf RM, DeMets DL. Principles from clinical trials relevant to clinical practice: Part II. *Circulation* 2002;106:1172-5.
27. Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Prog.Cardiovasc. Dis.* 1985;27:335-71.
28. Taylor DW, Sackett DL, Haynes RB. Sample size for randomized trials in stroke prevention. How many patients do we need? *Stroke* 1984;15:968-71.
29. Grimes DA, Schulz KF. Determining sample size and power in clinical trials: the forgotten essential. *Semin.Reprod.Endocrinol.* 1996;14:125-31.
30. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;310:452-4.
31. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2. ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. *Lancet* 1988;2:349-60.