

O que é um estudo clínico randomizado?

What is a randomized clinical trial?

Raphael F. de Souza¹

RESUMO

O estudo clínico randomizado (ECR) é uma das ferramentas mais poderosas para a obtenção de evidências para o cuidado à saúde. Apesar de algumas possíveis variações, baseiam-se na comparação entre duas ou mais intervenções, as quais são controladas pelos pesquisadores e aplicadas de forma aleatória em um grupo de participantes. O objetivo deste artigo foi descrever aspectos relativos à validade externa e interna dos ECRs, bem como apresentar bases de dados para sua obtenção e ferramentas para avaliação de qualidade.

Palavras-chave: Ensaio Clínico Controlado Aleatório. Ensaio Controlado Aleatório. Ensaio Clínico Controlado. Medicina Baseada em Evidências.

O que é um estudo clínico randomizado?

O estudo clínico randomizado (ECR) consiste basicamente em um tipo de estudo experimental, desenvolvido em seres humanos e que visa o conhecimento do efeito de intervenções em saúde. Pode ser considerado como uma das ferramentas mais poderosas para a obtenção de evidências para a prática clínica. Associada a esse poder, encontra-se a simplicidade em seu desenho, quando comparado a outros tipos de estudos¹. Bons ECRs são capazes de minimizar a influência de fatores de confusão sobre relações de causa-efeito, quando comparados aos demais desenhos, daí sua grande relevância como fonte de evidências². No caso específico do ECR, diferencia-se dos demais tipos de estudos clínicos experimentais pelo

fato de que os participantes recebem uma dentre as intervenções propostas de forma aleatória¹.

No contexto da classificação dos estudos científicos em saúde, pode-se dizer que os ECRs são, dentre os estudos primários, os de maior relevância para a clínica³. As revisões sistemáticas, apesar de serem consideradas ainda mais relevantes, são estudos secundários, ou seja, dependem de estudos primários com qualidade para derivarem inferências. Daí a grande importância dos ECRs como fonte de evidências também para as revisões sistemáticas⁴.

A fim de determinar se um artigo científico refere-se ao relato de um ECR, quatro características fundamentais devem estar presentes em um artigo científico. A primeira delas é a comparação entre intervenções aplicadas à saúde de seres humanos, ou seja, estudos sobre aspectos sem relação direta com

1. Docente da Faculdade de Odontologia de Ribeirão Preto da Universidade de São Paulo.

Correspondência:
Departamento de Materiais Dentários e Prótese
Faculdade de Odontologia de Ribeirão Preto - USP.
Avenida do Café, s/nº
14040-050 - Ribeirão Preto, São Paulo, Brasil.

Artigo recebido em 18/12/2008
Aprovado em 20/05/2009

a saúde ou desenvolvidos em animais são desconsiderados. Um segundo aspecto relevante é a natureza prospectiva; as intervenções são planejadas antes e a exposição é controlada pelos pesquisadores. Em terceiro lugar, duas ou mais intervenções são comparadas no estudo. Por intervenção, entendem-se, além de tratamentos para desordens físicas ou mentais, os regimes preventivos, programas de detecção ou testes diagnósticos (uma pode ser um grupo controle/sem tratamento). Além disso, uma ou mais das intervenções comparadas pode ser consistir em um grupo controle, sem nenhum procedimento ativo. Por fim, a aplicação das intervenções deve ser aleatória, tendo apenas a chance influenciando o processo⁵.

O desenho típico de um ECR pode ser observado na Figura 1. Uma primeira característica é o recrutamento de um grupo comum, a partir de uma população de interesse. Em seguida é que se decidem quais intervenções os participantes receberão, por meio da randomização. Tecnicamente, esse processo envolve determinar a alocação por meio de números obtidos por sorteio, ao invés de características da amostra ou preferência dos participantes. Após a aplicação das intervenções, realiza-se a leitura de uma ou mais variáveis de desfecho que, na figura, apresenta-se como sucesso e insucesso⁶.

Variações⁶

O desenho típico de ECR, também denominado ensaio randomizado de grupos paralelos, não é a única forma de apresentação. Algumas variações são comuns, e dentre elas encontra-se o uso de um período pré-teste, entre o recrutamento dos participantes e a randomização. Por exemplo, pode-se tratar todos os

participantes nesse estágio com um placebo ou tratamento ativo, com o intuito de detectar participantes com baixa complacência. Conforme será explicado na seção “Validade Interna”, a perda de participantes durante a aplicação das intervenções pode gerar um tipo de viés; portanto, seriam melhor que as perdas acontecessem em um período prévio à alocação dos participantes às intervenções.

Existe também a possibilidade de se aplicar um desenho fatorial, ou seja, dois ou mais fatores de variação dentro do mesmo ECR. Supondo que um pesquisador quer testar para uma determinada desordem muscular um procedimento fisioterápico vs. um procedimento inativo, um ECR seria facilmente delineado. No entanto, poderia ser interessante testar também em outro ECR o efeito de um antiinflamatório vs. placebo; nesse caso, o mesmo ECR poderia comportar ambos os fatores de variação, ou conjunto de intervenções. Assim, seria delineado um estudo onde quatro possibilidades existem, ou seja, o participante recebe:

- 1) o tratamento fisioterápico e o antiinflamatório;
- 2) o tratamento fisioterápico e o placebo;
- 3) o procedimento inativo e o antiinflamatório; ou
- 4) o procedimento inativo e o placebo.

Dessa forma, têm-se duas análises separadas, em função de cada fator, e analisa-se também a interação entre ambos.

Por fim, é possível testar-se duas ou mais intervenções usando-se o mesmo participante. Esse tipo de ECR é denominado cruzado (*cross-over*) e tem a vantagem de ter o mesmo indivíduo como controle dele mesmo, ou seja, emprega-se uma análise pareada ao invés de grupos independentes. A maior vantagem

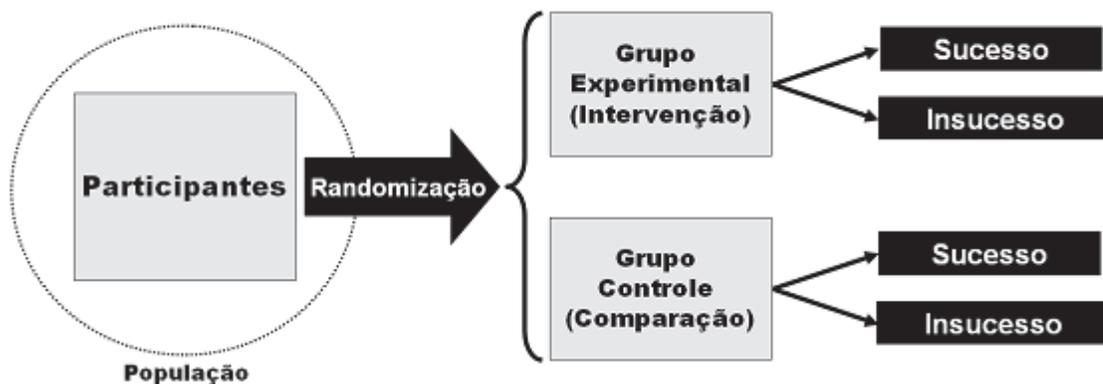


Figura 1. Desenho típico de um ECR (adaptado de Hulley et al.⁶).

desse processo é o aumento do poder estatístico, o que possibilita testar hipóteses com um número menor de participantes. Por outro lado, nem todas as intervenções são reversíveis; por exemplo, seria impossível conduzir um ECR cruzado para a maioria dos procedimentos cirúrgicos. Além disso, caso haja um efeito residual de uma intervenção durante a aplicação da próxima, teríamos um viés importante (efeito *carry-over*). Nesses casos, indica-se como norma o planejamento de um período de *washout* entre as intervenções, a fim de esperar a cessação do efeito residual. A Figura 2 contém um esquema de como se estrutura um ECR cruzado. Nos ECRs em se randomiza seqüências de 3 ou mais intervenções para cada participantes, usa-se o termo quadrado latino (em inglês, *Latin square*).

Validade Externa

Independente do desenho do ECR, algumas características devem ser observadas, a fim de extrapolar os resultados de um estudo para determinada realidade clínica. Essa possibilidade de generalizar os dados de pesquisa para uma circunstância cotidiana é denominada validade externa. Quanto mais as características consideradas pelo estudo forem próximas de determinada realidade de trabalho, mais os resultados do ECR podem ser extrapolados; ainda, mais relevante se torna a evidência dentro dessa realidade. Como um exemplo, seria improvável que amostras provenientes de centros altamente especializados possam ser representativas do que acontece em clínicas de atenção básica. Especificamente, quatro aspectos básicos determinam a validade externa:

- 1) participante;
- 2) intervenções;
- 3) ambiente;
- 4) desfechos⁷.

Com respeito aos participantes, é importante que as características sejam próximas àquelas da população de interesse. Aspectos como idade, gênero, severidade de doença e fatores de risco, co-morbidades associadas devem ser vistos com atenção⁷. Em ECRs, os participantes não são necessariamente pacientes, podem ser voluntários, parentes dos pacientes (cuidadores de idosos com doença de Alzheimer, por exemplo) ou aglomerados (randomiza-se grupos como postos de saúde ou escolas)¹. Intervenções podem apresentar diferentes regimes de aplicação, com distinção de dose, períodos e vias de administração. O uso de tratamento simultâneo ao ECR, bem como categorias dentro de certas modalidades, pode influenciar os resultados e devem também ser pesados⁷. Além disso, é de se esperar que diferentes ambientes de estudo possam acarretar diferentes resultados. A pergunta que se faz é o quanto o ambiente do estudo assemelha-se à clínica, pois diferença em fatores como o nível de atenção (primária a terciária) e experiência do profissional pode diminuir a validade externa⁷. Por fim, o tipo e definição das variáveis de desfecho são altamente importantes. Destaca-se o uso de variáveis de pouca relevância clínica, ao invés daquelas de real valor para o participante. É comum que variáveis secundárias mostrem maior eficácia para determinado tratamento, porém essa diferença não é detectável clinicamente. Por exemplo, do que adiantaria um procedimento capaz de aumentar a ejeção de sangue pelo coração após infarto do miocárdio, caso não haja melhoras na qualidade e expectativa de vida do pacien-

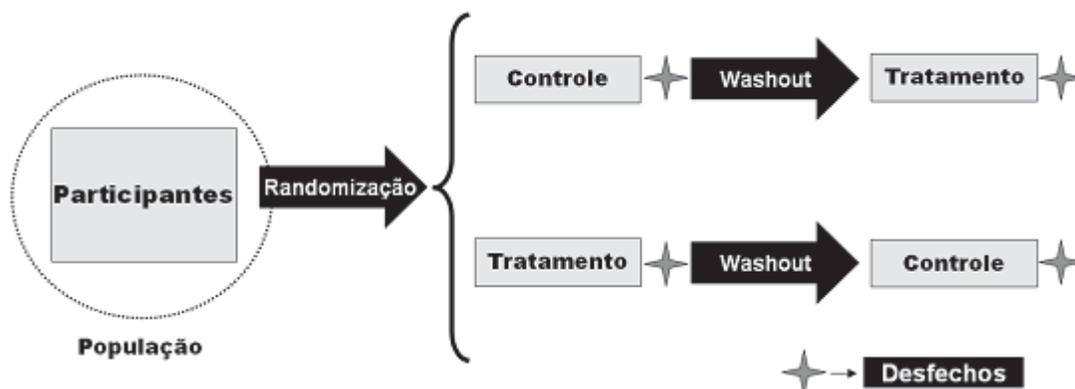


Figura 2. Desenho de um ECR cruzado (adaptado de Hulley et al.⁶).

te? Uma última característica relevante para os desfechos é que se considere o tempo de acompanhamento no ECR. Diferentes tempos podem apresentar diferentes resultados⁷.

Validade Interna

Questões referentes à validade interna direcionam-se ao quanto o erro sistemático (viés) é minimizado, independente das características descritas anteriormente. Quatro tipos principais de viés podem ser considerados em um ECR: seleção, desempenho, detecção e atrito. Cada um tem origem diferente e pode ser minimizado por manobras específicas, como a randomização⁷.

O viés de seleção simboliza a presença de alguma diferença sistemática entre os grupos comparados. O ideal seria que os participantes fossem alocados nos grupos sem nenhuma influência, e todo participante teria a mesma chance de ser alocado em qualquer um dos grupos ou sequência. Duas manobras são utilizadas para minimizar essa fonte de viés. A primeira delas, a randomização, implica em usar uma sequência de alocação que previna o viés de seleção. As intervenções são sorteadas para os participantes, usando-se programas de computador ou tabelas de números aleatórios, por exemplo. A segunda manobra, a alocação oculta, parte do princípio de que há a possibilidade de se rejeitar determinados participantes em certos grupos. Por exemplo, poderia se evitar a alocação de casos mais graves em grupos com uso de placebo. Assim, uma forma de se evitar essa rejeição seria ocultar a sequência de alocação antes de se aplicar as intervenções. Não é sinônimo de cegamento, porque ocorre em período mais precoce do estudo e visa minimizar outro tipo de viés. A alocação oculta em geral é possível por meio de um pesquisador sem contato com os participantes e com o trabalho dos demais⁷. Tecnicamente, emprega-se, um escritório ou farmácia central (fora do ambiente de pesquisa), recipientes pré-codificados serialmente, software capaz de alocar após entrada de participante ou envelopes opacos e selados⁴.

Por sua vez, o viés de desempenho consiste naquelas diferenças sistemáticas no cuidado fornecido para os participantes em função dos grupos, e pode se originar de diversas fontes. Contaminação (intervenção no grupo controle) e co-intervenção (cuidado adicional a algum dos grupos) são possíveis causas. Ainda, participantes podem relatar mais sintomas caso

saibam a que grupo pertence e avaliadores de desfecho podem realizar análises tendenciosas caso conheçam o tratamento em curso. O manejo do viés de desempenho se dá por meio do cegamento, que deve ser realizado sempre que possível nas diversas partes do estudo. Essas compreendem o pesquisador responsável por aplicar as intervenções, os participantes, avaliadores de desfecho e mesmo o responsável pela análise estatística.

Um terceiro tipo de viés é o de detecção, que se refere às diferenças sistemáticas entre grupos durante à análise dos desfechos. Também é minimizado pelo cegamento, neste caso especificamente dos avaliadores de desfecho⁴.

Por fim, outra forma de erro sistemático é representada pelo viés de atrito, que corresponde à diferença sistemática entre grupos devido à perda de participantes do estudo. É sinônimo de viés de exclusão, e pode representar problemas quando os participantes deixam de retornar após o início das intervenções devido a, por exemplo, efeitos colaterais decorrentes de uma das intervenções⁴. Pode ser manejado por meio de um número mínimo de perdas no acompanhamento; mais de 20% de perda em um estudo é forte indicativo de viés, bem como perdas desiguais nos diferentes grupos. Assim, é fundamental que os autores descrevam se houve perdas em um ECR e os motivos. Além disso, cada participante, mesmo que não receba a intervenção alocada, deve ser analisado ao fim do estudo como se tivesse recebido. A esse princípio designa-se intenção-de-tratar.⁷

Bases de Dados

É possível conseguir referências de ECRs por meio de bases de dados internacionais como o Medline, EMBASE, ISI, entre outras. Bases regionais, como a LILACS, podem e devem ser empregadas, principalmente quando características regionais podem influenciar na validade externa. Independente do caso, uma busca direta usando termos relativos aos participantes, intervenções e desfechos de interesse poderiam ser usados na base eletrônica; porém, a adição de termos relativos ao desenho torna a busca mais ágil. As bases de dados em saúde podem fornecer descritores referentes aos ECRs, como o termo “randomized controlled trial [pt]” do MeSH (Medline). Buscas com maior sensibilidade podem ser obtidas usando filtros de busca, especialmente quando se trata de obter ECRs para revisões sistemáticas, Quadro 1⁴.

Quadro 1

Estratégia de busca de alta sensibilidade para identificação de ECR no MEDLINE (formato PubMed)

#1	randomized controlled trial [pt]
#2	controlled clinical trial [pt]
#3	randomized [tiab]
#4	placebo [tiab]
#5	drug therapy [sh]
#6	randomly [tiab]
#7	trial [tiab]
#8	groups [tiab]
#9	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8
#10	animals [mh] not (humans [mh] and animals [mh])
#11	#9 not #10

Outra fonte importante para a obtenção de registros de ECRs é a Cochrane Central Register of Controlled Trials (CENTRAL), da Biblioteca Cochrane. A CENTRAL apresenta uma lista de resumos de ECRs e estudos clínicos controlados, que foram obtidos por meio de busca manual de periódicos em diversas áreas da saúde.

Ainda, mesmo outras fontes como bases de dissertações e teses, resumos de conferências podem fornecer informações úteis para a prática clínica. Dados ainda não publicados em periódicos podem ser conferidos com antecedência. Por fim, as revistas de saúde baseada em evidências também são fontes importantes, pois apresentam resumos críticos da literatura corrente. Em geral, sua vantagem é facilitar para o clínico a visualização de possíveis limitações de estudos publicados recentemente.

Avaliação de Qualidade

Por fim, após a obtenção de ECRs a partir das possíveis bases, seria interessante que o clínico conseguisse identificar limitações importantes. Características associadas à validade interna e externa podem ser avaliadas sistematicamente por meio do conhecimento de possíveis vieses. Essa avaliação, contudo, pode ser facilitada por meio de instrumentos específicos. Primeiramente, destaca-se o CONSORT, por listar uma série de características que deveriam

ser descritas em um ECR⁸. Destaca-se também o recente método GRADE (*Grading of Recommendations Assessment, Development and Evaluation*), por classificar estudos de diferentes desenhos de acordo com a qualidade e relevância da evidência, separadamente⁹.

Com base nos quatro modos de controle de vieses, é possível classificar a qualidade dos ECRs. Algumas revisões sistemáticas da Colaboração Cochrane têm apresentado um instrumento de análise de qualidade com essas características, o qual pode ser conferido abaixo:

1) Randomização

- Adequada:** um dos métodos a seguir foi descrito pelo artigo: números gerados por computador ou tabela de números aleatórios, cara ou coroa, sorteios de cartas ou jogo de dados.
- Não descrita;
- Inadequada:** sequência gerada por meio de número de prontuário, data de nascimento ou números alternados.

2) Alocação Oculta

- Adequada:** alocação gerada centralmente ou por meio de envelopes opacos, selados e numerados sequencialmente.
- Não descrita;
- Inadequada:** houve alocação aberta e se os participantes ou pesquisadores poderiam prever a alocação.

3) Cegamento

Se houve cegamento (ou não) para as seguintes partes:

- Participantes (sim/ não / não descrito / não se aplica);
- Pesquisador (sim/ não / não descrito / não se aplica);
- avaliador de desfechos (sim/ não / não descrito / não se aplica);
- analista dos dados (sim/ não / não descrito / não se aplica).

4) Manejo de perdas e desistências

- Sim:** se há uma descrição clara sobre as diferenças entre dois ou mais grupos testados durante o acompanhamento;
- Não descrito;
- Não.

Após a análise, é possível classificar os ECRs como sujeitos a diferentes riscos de vies:

- a) Baixo risco (possível viés com improvável influência sobre os resultados) se todos os critérios forem adequados;
- b) Risco moderado (possível viés que levanta alguma dúvida referente aos resultados) se um ou mais critérios são parcialmente adequados.
- c) Alto risco (possível viés que enfraquece seriamente a confiança nos resultados) se um ou mais critérios são inadequados.

ABSTRACT

The randomized clinical trial (RCT) is one of the most powerful tools for the obtainment of evidence for healthcare. Despite some possible variations, they are based on the comparison between two or more interventions, which are controlled by researchers and randomly applied on a group of participants. The aim of this study was to describe aspects of external and internal validity of RCTs, as well as to present databases for their obtainment and quality assessment tools.

Key words: Randomized Controlled Trial. Controlled Clinical Trial. Evidence-Based Medicine.

Referências

1. Jadad AR, Enkin MW. Randomized controlled trials. Questions, answers, and musings. 2nd. ed. London: Blackwell Publishing/BMJ Books, 2007.
2. Clancy MJ. Overview of research designs. *Emerg. med. j.* 2002; 19: 546-9.
3. Albrektsson T, Wennerberg A. Oral implant surfaces: Part 2 – review focusing on clinical knowledge of different surfaces. *Int. j. prosthodont.* 2004; 17: 544-64.
4. Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester: John Wiley & Sons; 2008. 3.
5. Bickley SR. Oral Health Review Group Journal Handsearchers' Manual. Manchester: Cochrane Oral Health Group; 2002.
6. Hulley SB, Cummings SR, Browner WS, Grady D, Hearst N, Newman TB. *Delineando a pesquisa clínica: uma abordagem epidemiológica*. 2nd. ed. Porto Alegre: Artmed; 2003.
7. Jüni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *BMJ* 2001; 323: 42-6.
8. Altman DG. Better reporting of randomised controlled trials: the CONSORT statement. *BMJ* 1996; 313: 570–1.
9. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008; 336: 924-6.