

Avaliação Crítica de um Ensaio Clínico Aleatorizado (Parte II): Do Seguimento dos Participantes à Análise de Resultados

Critical Evaluation of a Randomized Clinical Trial (Part II): From the Follow-up of Participants to the Analysis of Results

Bernardo Sousa-Pinto^{1,2}, Luís Azevedo^{1,2,3*}

Autores

¹ MEDCIDS - Departamento de Medicina da Comunidade, Informação e Decisão em Saúde, Faculdade de Medicina da Universidade do Porto, Universidade do Porto, Porto, Portugal.

² CINTESIS - Centro de Investigação em Tecnologias e Serviços de Saúde, Porto, Portugal.

³ Centro Nacional de Observação em Dor, Faculdade de Medicina da Universidade do Porto, Porto, Portugal.

Palavras-chave

Investigação Biomédica; Medicina Baseada em Evidências; Tomada de Decisões

Keywords

Biomedical Research; Decision Making; Evidence-Based Medicine

INTRODUÇÃO

Esta constitui a terceira publicação de uma série da Revista da Sociedade Portuguesa de Anestesiologia dedicada à discussão de aspectos metodológicos e operacionais da investigação clínica. O artigo anterior desta série iniciou uma discussão relativa à avaliação crítica de ensaios clínicos aleatorizados (ECAs), versando dois aspectos fundamentais nesse contexto - os processos de selecção dos participantes e de atribuição aleatória da intervenção.¹ Nesse sentido, o presente artigo tem por objectivo dar continuidade ao anterior, focando-se em dois outros aspectos que são, também eles, fundamentais e devem ser sempre tidos em conta no decurso da avaliação crítica de um ECA - o seguimento dos participantes e os métodos de análise de resultados.

À semelhança do que se verificou anteriormente, a abordagem conceptual subjacente a este artigo tem por base os diferentes itens da secção de Métodos das *guidelines* CONSORT (fundamentais para a redacção, mas também para a análise crítica e estruturada de um ECA),² e é acompanhada por um exemplo prático, correspondendo à análise do ECA cuja avaliação crítica se iniciou no artigo anterior (Weinberg L, *et al.* *A randomised controlled trial of peri-operative lidocaine infusions for open radical prostatectomy.* *Anaesthesia.*

2016;71:405-410 - neste estudo, os autores procuram avaliar o efeito da administração peri-operatória de lidocaína 2% em homens submetidos a prostatectomia retro-púbica).³

SEGUIMENTO DOS PARTICIPANTES: OCULTAÇÃO E OUTROS CUIDADOS A TER NO FOLLOW-UP

Tendo os participantes sido alocados aos respectivos grupos, estes passam a ser seguidos durante um determinado período de seguimento (“*follow-up*”) que se deseja adequado face às variáveis de resultado definidas - por exemplo, o seguimento de um mês seria inadequado para avaliar os efeitos de um bifosfonato no risco de fracturas ósseas.

Durante o período de *follow-up*, é fundamental assegurar que o seguimento é o mais completo possível para todos os participantes - devem ser minimizadas as desistências e, sempre que estas ocorrem, procurar registar o seu motivo. Ao analisar um ECA, é importante atentar não só na percentagem de perdas de seguimento, mas até que ponto estas são diferenciais entre os vários grupos em comparação. Ao longo do período de *follow-up*, deverão ainda verificar-se estratégias que assegurem que todos os participantes são alvo de um igual seguimento, independentemente do grupo ao qual foram alocados - uma das estratégias mais importantes com esse objectivo consiste na ocultação.

OCULTAÇÃO

O processo de ocultação (“*blinding*”) consiste no

*Autor Correspondente:

Luís Filipe Azevedo

Morada: MEDCIDS – Departamento de Medicina da Comunidade, Informação e Decisão em Saúde, Faculdade de Medicina da Universidade do Porto, Rua Dr. Plácido da Costa, 4200-450, Porto, Portugal.

E-mail: lazevedo@med.up.pt

desconhecimento do grupo ao qual os participantes de um ECA foram alocados.⁴ A ocultação pode ser feita a vários níveis, envolvendo diferentes intervenientes do estudo - verifica-se uma ocultação simples quando apenas os participantes desconhecem o grupo ao qual foram alocados; já a ocultação dupla implica desconhecimento tanto por parte dos participantes como dos investigadores. Desejavelmente, a ocultação deverá também estender-se a outros níveis, nomeadamente aos responsáveis pela medição do *outcome* e pela análise estatística.⁵ A ocultação tem por objectivo minorar o risco de vieses de informação - de facto, podem verificar-se diferenças (ainda que inconscientes) de seguimento, tratamento ou medição do *outcome* caso os investigadores conheçam o grupo ao qual foram alocados os participantes; por outro lado, também os participantes podem evidenciar comportamentos diferentes caso saibam se estão a ser alvo da intervenção experimental ou do seu comparador.⁴ Em suma, no decurso de um ECA, é fundamental assegurar idêntico *follow-up* a todos os participantes independentemente do grupo ao qual tenham sido alocados, sendo que a ocultação constitui um processo que a tal auxilia.

Para que a ocultação seja eficaz e para controlar o efeito placebo (efeito inespecífico de melhoria simplesmente resultante do facto de o participante achar que está a ser alvo de uma intervenção terapêutica⁶), é fundamental providenciar a administração de um placebo (ou outro comparador activo) ao grupo que não é alvo da intervenção experimental. O placebo consiste numa intervenção sem efeitos aparentes na variável de resultado em estudo, mas que deve mimetizar o mais possível a intervenção activa - o exemplo mais imediato é pensar num comprimido *dummy* com a mesma cor, tamanho e forma face ao que está a ser testado⁶ (note-se que, mesmo que a intervenção experimental esteja a ser comparada com uma alternativa *standard*, deve ser providenciado placebo aos participantes - nesses casos, o grupo controlo deverá receber a alternativa *standard* + placebo; por sua vez, o grupo experimental deverá receber a alternativa *standard* + intervenção experimental). Este ponto é de tal forma relevante que quando estão a ser avaliados procedimentos cirúrgicos, pode optar-se por um procedimento simulado (*sham*), em que o participante é levado ao bloco operatório e anestesiado (podendo até ser alvo de uma incisão) mas não é alvo da cirurgia em estudo.⁷ Embora nestes casos tenham que ser ponderadas e pesadas da forma mais equilibrada possível as questões éticas e metodológicas, as recomendações existentes apontam neste sentido. Por outro lado, quando estão a ser testados dispositivos médicos, podem ser aplicados dispositivos simulados ou desligados. Não obstante, nem sempre é possível proceder a ocultação, particularmente com alguns tipos de intervenção (e.g., aplicações de telemóvel ou intervenções educacionais). Assim, não é obrigatório que ocorra ocultação num ECA, embora a sua não-execução

potencie o risco de vieses de informação (num grau variável consoante o que está a ser avaliado - a ausência de ocultação do responsável pela avaliação do *outcome* poderá não constituir um grande problema caso se ambicione medir um “*outcome* duro” como a morte).

Procedendo-se a ocultação num ECA, é fundamental nomear a que a nível e de que modo esta foi feita. No estudo de Weinberg *et al.*,³ verificou-se ocultação dos participantes e profissionais de saúde que com eles contactavam - um farmacêutico clínico independente era responsável por preparar infusões semelhantes que podiam conter lidocaína (intervenção experimental) ou uma solução salina (comparador). Seria desejável uma confirmação que a ocultação se estendeu à avaliação dos *outcomes* - pese embora o facto de se verificar um *outcome* primário “objectivo” (tempo de internamento), poderá existir alguma subjectividade na aplicação dos critérios para alta.

ANÁLISE DOS RESULTADOS CONSIDERAÇÕES SOBRE A DEFINIÇÃO DAS VARIÁVEIS DE RESULTADO

Um ECA pode contemplar uma ou mais variáveis de resultado (*outcome*). Quando existe mais do que um *outcome*, o *outcome* primário corresponde à variável de resultado com base na qual é calculado o tamanho amostral.⁸ Neste sentido, nem sempre o *outcome* primário constitui a variável clinicamente mais relevante. Este fenómeno pode ser compreensível sob o ponto de vista da viabilidade do estudo, particularmente quando os *outcomes* mais relevantes são relativamente raros ou ocorrem a muito longo prazo, exigindo o recrutamento de um grande número de participantes ou o seu seguimento durante um extenso período de tempo. O reverso da medalha é a determinação do tamanho amostral com base em variáveis que não são as de maior interesse clínico.

Como forma de tentar ultrapassar esta limitação, é possível definir *endpoints* compostos - ou seja, agregar vários *outcomes* (e.g., morte, agudização, internamento...) numa única variável de resultado, bastando o participante desenvolver um destes eventos para se considerar a ocorrência dessa mesma variável.⁸ A definição de *endpoints* compostos deve ser alvo de um atento escrutínio, uma vez que estes podem incluir *outcomes* comuns mas relativamente pouco relevantes (e.g., desenvolvimento de dor torácica num *endpoint* composto de avaliação de eventos cardiovasculares).

Uma outra estratégia frequentemente utilizada passa pela definição de *endpoints* surrogados - variáveis ou biomarcadores que, por se relacionarem com um dado *outcome* clínico, são avaliados em detrimento deste último.^{8,9} Um exemplo clássico de um *endpoint* surrogado prende-se com a avaliação da variação da densidade mineral óssea como biomarcador do risco de fracturas ósseas. Assim, perante a definição de *endpoints* surrogados, importa perceber até que

ponto existe evidência que efeitos nesses biomarcadores se associam de facto a um efeito real nas variáveis de interesse clínico, pois nem sempre essa relação é clara e devidamente demonstrada.

Adicionalmente, importa aferir de que forma foram avaliadas as variáveis de resultado - quais os métodos e instrumentos utilizados, quando se procedeu a essa avaliação, quem foram os elementos responsáveis pela medição do *outcome*, e até que ponto existia ocultação do grupo ao qual tinha sido alocado cada participante. Caso as variáveis de resultado sejam alteradas após início do ECA, tal deverá ser reportado e devidamente justificado.

No estudo de Weinberg *et al.*,³ o *outcome* primário consistiu no tempo de internamento, o qual constitui uma variável de resultado relevante per se, não consistindo num *outcome* composto nem num *endpoint* surrogado. Por outro lado, os *outcomes* secundários avaliados incluíram a dor pós-operatória, consumo de analgésicos, desenvolvimento de eventos adversos e satisfação dos participantes. O modo como cada *outcome* foi avaliado encontra-se sumariamente descrito, pese embora os instrumentos e o *timing* da medição não se encontrarem descritos para todas as variáveis (e.g., satisfação dos participantes). Não é providenciada informação relativa aos elementos responsáveis pela medição dos *outcomes*.

ANÁLISE DOS RESULTADOS FACE AO OBJECTIVO DO ECA

Antes de se discutir alguns aspectos mais directamente relacionados com a análise dos resultados, interessa frisar que, nos ECA, revela-se fundamental proceder ao cálculo do tamanho amostral adequado. Caso contrário, eventuais diferenças clinicamente relevantes podem não ser observadas simplesmente pelo facto de o estudo não ter um número suficiente de participantes que assegure uma precisão das estimativas capaz de as detectar.¹⁰ A existência de um tamanho amostral adequado é, assim, essencial para diminuir o impacto de eventuais erros aleatórios.

Um outro aspecto importante a ter em conta prende-se com o objectivo com o qual o ECA foi concebido - nem todos os estudos são conduzidos com o objectivo de demonstrar superioridade de uma dada intervenção face ao seu comparador; o objectivo do ECA pode simplesmente passar por avaliar se a intervenção experimental é equivalente ou não-inferior a um dado comparador.¹¹ Por exemplo, imagine-se que existe uma terapia standard que, sendo bastante eficaz, comporta um elevado risco de eventos adversos - neste contexto, mais do que uma nova terapia mais eficaz, importaria desenvolver uma alternativa que se revelasse mais segura, desde que tão eficaz (ou que não menos eficaz) que a terapia *standard*. Nesse sentido, os resultados de um ECA devem ser apresentados tendo em conta o objectivo com que o estudo foi concebido - não podem ser tomadas conclusões

de superioridade num ECA desenhado para equivalência ou não-inferioridade, tal como não podem ser tomadas conclusões de equivalência ou não-inferioridade num estudo desenhado para demonstrar superioridade. De facto, o cálculo do tamanho amostral e os métodos de análise são diferentes consoante o ensaio seja desenhado para demonstrar superioridade, equivalência ou não-inferioridade.¹¹

No estudo de Weinberg *et al.*,³ o cálculo do tamanho amostral é descrito no último parágrafo dos Métodos, tendo sido feito com objectivo de detectar uma redução média de 18 horas no tempo de internamento. Assim, embora tal não esteja explícito, pressupõe-se que o ECA foi levado a cabo com o objectivo de demonstrar superioridade.

ANÁLISE POR INTENÇÃO DE TRATAR VERSUS ANÁLISE POR PROTOCOLO

Numa situação ideal, todos os participantes manter-se-iam no estudo até ao final do período de seguimento, permanecendo sempre no grupo ao qual tinham sido aleatoriamente alocados. Infelizmente, tal não se verifica em praticamente nenhum ECA - tipicamente, existem participantes que desistem de participar no estudo, são impedidos de completar todo o período de seguimento, ou acabam por receber a intervenção de um grupo diferente daquele ao qual tinham sido inicialmente alocados.

Nesse sentido, existem duas possibilidades no que respeita ao modo como se procede à análise dos resultados de um ECA. Numa análise por protocolo, os participantes são analisados de acordo com a intervenção que completaram no final do estudo. Por contraponto, numa análise de acordo com a intenção de tratar, os participantes são avaliados tal como foram alocados pela atribuição aleatória da intervenção, mesmo que subsequentemente tenham “mudado de grupo” ou desistido.¹² Embora possa inicialmente parecer contra-intuitivo, esta última forma de análise deve ser aquela que deverá ser preferencialmente reportada num ECA. De facto, enquanto a análise por protocolo apenas informa acerca da eficácia da intervenção experimental num contexto ideal (porque limita a análise ao grupo que a completou), a análise por intenção de tratar informa acerca da efectividade da mesma; ou seja, espelha o desempenho da intervenção num contexto mais próximo do da prática clínica real.¹² A título de exemplo, imagine-se que está a ser testada uma intervenção que, sendo bastante eficaz, se demonstra bastante mal tolerada - uma análise por protocolo iria sobrestimar os benefícios dessa intervenção, pois excluiria os (muitos) participantes que não a tinham conseguido tolerar; por sua vez, a análise por intenção de tratar constataria que, dado ser mal tolerada, a intervenção poderia não se revelar mais efectiva na prática clínica que o seu comparador. Além disto, e porque as razões de perda ou mudança de grupo estão muitas vezes associadas aos resultados clínicos observados, é possível demonstrar que

a análise segundo a intenção de tratar é aquela que melhor aproxima o real efeito causal da intervenção em estudo e a que permite um mais adequado controlo das várias fontes de enviesamento possíveis.

No estudo de Weinberg *et al.*,³ não se encontra explícito o tipo de análise realizada. De qualquer forma, é possível verificar que, não tendo ocorrido desistências de participantes nem “mudanças de grupo”, um participante do grupo experimental que recebeu uma dose errada de lidocaína acabou por ser incluído na análise.

ANÁLISE DE SUBGRUPOS

Por vezes, para além da análise principal, são conduzidas análises *a posteriori* para subgrupos específicos de participantes. Ou seja, os investigadores testam num conjunto restrito de indivíduos (e.g., apenas pacientes > 65 anos) se existem diferenças entre os participantes alocados aos diferentes grupos.¹³ Esta estratégia é frequentemente conduzida quando não se observam diferenças estatisticamente significativas no respeitante à análise principal.

Todavia, não tendo sido conduzida uma atribuição aleatória da intervenção para cada subgrupo em análise, a evidência resultante das análises de subgrupos deve ser vista como resultados meramente exploratórios e que carecem de confirmação posterior. Por outras palavras, sendo feitas *a posteriori*, as análises de subgrupos acabam por ser meramente geradoras de hipóteses, pois não temos as mesmas garantias de comparabilidade dos grupos nem de poder adequado para a observação dos efeitos pretendidos. Adicionalmente, as análises de subgrupos colocam vários problemas estatísticos e metodológicos.¹³ Nesse sentido, as análises de subgrupos diferem dos estudos com aleatorização estratificada, nos quais a atribuição aleatória da intervenção é feita dentro de cada estrato na fase inicial de desenho do estudo, permitindo dar garantias adequadas sobre o poder e a comparabilidade dos grupos. Assim, importa perceber que as análises de grupos específicos de participantes são adequadas se surgirem no contexto de uma aleatorização estratificada *a priori*, mas têm que ser vistas com muito cuidado quando são definidas *a posteriori*, na fase de análise estatística do ECA, pois podem resultar em meros artefactos estatísticos e aleatórios.

No estudo de Weinberg *et al.*,³ não foram conduzidas análises de subgrupos. Uma possível análise de subgrupos no contexto deste estudo passaria pela comparação dos *outcomes* por hospital.

CONCLUSÃO

No que diz respeito à efectividade de intervenções e à determinação de relações de causalidade, a evidência providenciada pelos ECA é considerada superior à de outros

estudos. Tal deve-se ao facto de a correcta execução de um ECA compreender um conjunto de processos que actuam num sentido de minimizar o risco de diversas vieses de selecção, informação e confundimento. No entanto, não é suficiente que um estudo seja designado de ECA para garantir automaticamente a sua qualidade e validade. Na verdade, a validade e qualidade destes estudos dependerão sempre da aplicação mais ou menos adequada e completa dos preceitos metodológicos atrás enunciados. Desta forma, é fundamental proceder a uma apreciação crítica relativamente ao modo como esses mesmos processos são levados a cabo - ao longo destas duas publicações, discutimos a importância de olhar criticamente para aspectos como os critérios de elegibilidade dos participantes, a atribuição aleatória da intervenção, a ocultação da sequência de alocação, a adequação do comparador, o processo de ocultação pós-aleatorização, a definição dos *outcomes* e o modo como os resultados são analisados. Adicionalmente, outros aspectos merecem a atenção dos leitores, incluindo a existência de conflitos de interesse e a adequação da discussão dos resultados. De ressaltar apenas que, pese embora a sua importância, nem sempre os ECA constituem a melhor opção para dar resposta a questões de investigação, sendo a sua execução frequentemente limitada por questões de índole ética e/ou prática. Para esses casos, torna-se necessário recorrer a outras tipologias de estudos, pese embora o seu maior potencial para alguns tipos de vieses - os estudos observacionais.

Este artigo foi escrito de acordo com a antiga ortografia, não estando ao abrigo do novo Acordo Ortográfico.

Responsabilidades Éticas

Conflitos de interesse: Os autores declaram não possuir conflitos de interesse.

Suporte Financeiro: O presente trabalho não foi suportado por nenhum subsídio ou bolsa.

Ethical Disclosures

Conflicts of interest: The authors have no conflicts of interest to declare.

Financing support: This work has not received any contribution, grant or scholarship.

Submissão: 30 de Agosto, 2018 | Aceitação: 20 de outubro, 2018
Received: 30th of August, 2018 | Accepted: 20th of October, 2018

REFERÊNCIAS

1. Sousa-Pinto B, Azevedo LF. Avaliação crítica de um ensaio clínico aleatorizado: Da selecção dos participantes à atribuição aleatória da intervenção. *Rev Soc Port Anesthesiol*. 2018; 27: 38-41.
2. Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomized trials. *BMJ*. 2010;340:698-702. doi: 10.1136/bmj.c332.
3. Karanicolos PJ, Farrokhhyar F, Bhandari M. Blinding: Who, what, when, why, how? *Can J Surg*. 2010;53:345-8.
4. Kim J, Shin W. How to do random allocation (randomization). *Clin Orthop Surg*. 2014;6:103-9. doi: 10.4055/cios.2014.6.1.103.
5. Papakostas YG, Daras MD. Placebos, placebo effect, and the response to the healing situation: The evolution of a concept. *Epilepsia*. 2001;42:1614-25.
6. Wolf BR, Buckwalter JA. Randomized surgical trials and “sham” surgery: Relevance to modern orthopaedics and minimally invasive surgery. *Iowa Orthop J*. 2006;26:107-11.
7. Weinberg L, Rachbuch C, Ting S, Howard W, Yeomans M, Gordon I, et al. A randomised controlled trial of peri-operative lidocaine infusions for open radical prostatectomy. *Anaesthesia*. 2016;71:405-10. doi: 10.1111/anae.13368.
8. Goldberg R, Gore JM, Barton B, Gurwitz J. Individual and composite study

- endpoints: Separating the wheat from the chaff. *Am J Med.* 2014;127:379-84. doi: 10.1016/j.amjmed.2014.01.011.
9. Aronson JK. Biomarkers and surrogate endpoints. *Br J Clin Pharmacol.* 2005;59:491-4.
 10. Charles P, Giraudeau B, Dechartres A, Baron G, Ravaud P. Reporting of sample size calculation in randomized controlled trials: review. *BMJ.* 2009;338:b1732.
 11. Lesaffre E. Superiority, equivalence, and non-inferiority trials. *Bulletin of the NYU Hosp Joint Dis.* 2008;66:150-4.
 12. Sedgwick P. Intention to treat analysis versus per protocol analysis of trial data. *BMJ.* 2015;350:h681. doi: 10.1136/bmj.h681.
 13. Naggara O, Raymond J, Guilbert F, Altman DG. The problem of subgroup analyses: An example from a trial on ruptured intracranial aneurysms. *AJNR Am J Neuroradiol.* 2011;32:633-6. doi: 10.3174/ajnr.A2442.