

[Home](#)[Conteúdo](#)jameslindlibrary.org

Experimentos controlados de tratamentos na atenção à saúde

Nos últimos cinquenta anos, o atendimento à saúde teve um enorme impacto sobre a possibilidade das pessoas viverem mais e ficarem livres de problemas sérios de saúde. Estima-se que o atendimento à saúde tem sido responsável por mais ou menos um terço e meio do aumento na expectativa de vida e uma média de cinco anos a mais livre de problemas crônicos de saúde (Bunker et al. 1994). Todavia, o público poderia ter obtido - e ainda pode obter - um benefício muito maior dos recursos mais importantes investidos em pesquisas voltadas para a melhoria da saúde. Além disso, alguns tratamentos mal-sucedidos do passado poderiam ter sido evitados e outros poderão ser evitados no futuro.

A *James Lind Library* foi criada para melhorar a compreensão geral sobre os experimentos controlados de tratamentos no atendimento à saúde e como eles evoluíram ao longo do tempo.

Declarações falsas sobre os efeitos de tratamentos são comuns, portanto, todos nós devemos entender como são produzidas as afirmações embasadas sobre os efeitos de tratamentos. Sem este conhecimento, corremos o risco de concluir que tratamentos ineficazes são eficazes, ou de que tratamentos eficazes são ineficazes.

Os experimentos controlados de tratamentos são testes que visam obter informações seguras sobre os efeitos de tratamentos, através da redução das influências equivocadas de [vieses](#) ou de obras do [acaso](#). Quando ignoramos a necessidade de experimentos controlados de tratamentos, as pessoas sofrem e morrem desnecessariamente.

Os artigos explicativos da *James Lind Library* foram escritos para promover um entendimento maior do motivo pelo qual os experimentos controlados de tratamentos são necessários, e no que eles consistem. Para acessar os artigos, clique nas palavras sublinhadas, abaixo, ou selecione-os na tela Contents ([Conteúdo](#)). Para fazer o download de todos os ensaios, de modo a serem impressos para serem lidos fora da tela, [clique aqui](#).

Os [experimentos controlados são necessários](#) porque existem muitas pessoas sendo prejudicadas inadvertidamente quando as decisões de tratamento não levam em conta evidências confiáveis.

Os [princípios dos experimentos controlados](#) têm evoluído há pelo menos um milênio - e ainda hoje continuam evoluindo.

As [comparações](#) são fundamentais para tratar de [incertezas legítimas](#) sobre os efeitos de tratamentos. As comparações controladas de tratamentos devem evitar [vieses](#), seja das [diferenças entre as pessoas comparadas](#) ou das [diferenças na forma como os desfechos dos tratamentos são avaliados](#). A identificação segura dos efeitos [imprevistos de tratamentos](#) apresenta desafios específicos.

A [interpretação de comparações imparciais](#) nem sempre é clara. Não raro, os efeitos do tratamento são desconsiderados devido às [diferenças entre os tratamentos planejados e os tratamentos recebidos](#). O [efeito do acaso](#) também pode levar a equívocos.

Os experimentos controlados de tratamentos devem levar em consideração todas as evidências relevantes. A preparação de avaliações sistemáticas de todas as evidências relevantes envolve a redução do impacto do [viés de informações e de seleção a partir da evidência disponível](#). Um método estatístico denominado [meta-análise](#) pode ajudar a evitar equívocos [causados pelo acaso](#) nas avaliações sistemáticas.

É preciso fazer [avaliações sistemáticas atualizadas de todas as evidências relevantes e confiáveis](#) dos experimentos controlados de tratamentos no atendimento à saúde. Mesmo com avaliações sistemáticas atualizadas, entretanto, é importante estar alerta quanto aos vieses e às "interpretações pessoais". Isto pode gerar avaliações distintas, que supostamente abordam a mesma questão, mas que chegam a conclusões contraditórias.

Em suma, a *James Lind Library* contém os seguintes artigos:

[Por que os experimentos controlados são necessários ?](#)

[Por que as comparações são fundamentais?](#)

[Por que as comparações devem tratar de dúvidas legítimas?](#)

Precaução nas comparações com vícios

Diferenças entre as pessoas comparadas

Diferenças na maneira como os resultados dos tratamentos são avaliados

Interpretação de comparações sem vícios

Diferenças entre tratamentos planejados e tratamentos de fato recebidos

Consideração do efeito do acaso

Identificação dos efeitos imprevistos do tratamento

Revisões sistemáticas das evidências relevantes

Como lidar com o vício da evidência disponível

Precaução com o vício da evidência disponível

Avaliações sistemáticas atualizadas das evidências relevantes e confiáveis

Avaliações sistemáticas atualizadas de todas as evidências relevantes e confiáveis

Estes artigos explicativos valeram-se de um rico material ilustrativo da James Lind Library. Para acessá-los, clique nos links sublinhados ou nas imagens dos arquivos.

O texto destes artigos pode ser copiado e usado para fins não comerciais, desde que seja feito com reconhecimento explícito à '**The James Lind Library (www.jameslindlibrary.org)**'.

O material contido nos artigos também fora incluído na obra de Evans, Thornton and Chalmers, '**Testing Treatments: better research for better health care**' - um livro de 100 páginas publicado em 2006 pela British Library. ISBN 0-7123-4909-X.

TESTING TREATMENTS BETTER RESEARCH FOR BETTER HEALTHCARE

How do we know whether a particular drug, therapy or operation really works, and how well? How reliable is the clinical evidence? Are clinical trials truly unbiased? And is current research fully focussed on the real needs of patients? Such timely and pressing questions are raised and resolved in this probing inquiry into modern clinical research, with far-reaching implications for daily medical practice and patient care. What emerges is the surprising truth that clinical research is neither as unbiased, nor as relevant as patients have every right to expect, but that everyone – patients, doctors and researchers – can do much to change current practice and achieve better healthcare.

Aimed at both patients and professionals, Testing Treatments builds a lively and thought-provoking argument for better, more reliable, more relevant research, with unbiased or 'fair' trials, and explains how patients can work with doctors to achieve this vital goal.

Expertly and thoroughly researched, but never dry or dull, the fast-moving commentary, spanning the gamut of illness and therapy – from mastectomy to thalidomide – explores a vast range of revealing case-studies, enlivened throughout by entertaining anecdotes and vivid eyewitness accounts drawn from the direct experience of patients, practitioners and researchers.

E, V, I, D, E, N, C, E

The
British Library

Cover image & design by
TRIO CHALMERS
www.triofolo.com

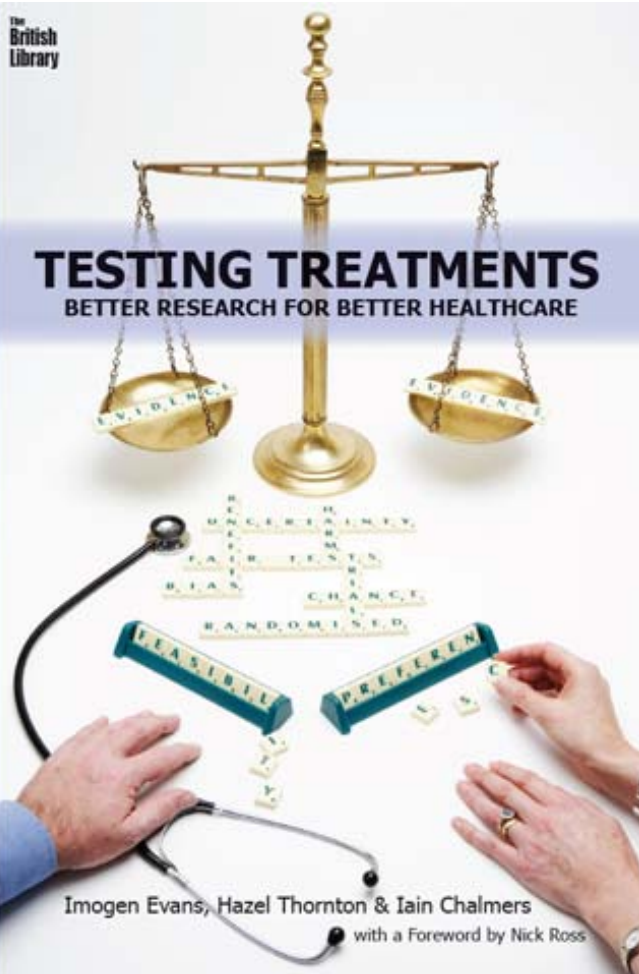
The
British
Library

TESTING TREATMENTS

Evans, Thornton & Chalmers

The
British
Library

TESTING TREATMENTS BETTER RESEARCH FOR BETTER HEALTHCARE



Cite as:

Editorial commentary (2007). Experimentos controlados de tratamentos na atenção à saúde. The James Lind Library (www.jameslindlibrary.org).

Next essay:

[Por que os experimentos controlados são necessários?](#)

Select essay:

Home

Conteúdo

[Envie comentários](#)

Por que os experimentos controlados são necessários?

Tentando trazer mais benefícios do que danos

Por que precisamos de experimentos controlados de tratamentos no atendimento à saúde? Os médicos não têm feito, há séculos, o "melhor" para os seus pacientes? Infelizmente, há muitos casos de médicos e outros profissionais da saúde que prejudicam seus pacientes porque as decisões de tratamentos não foram instruídas pelo o que agora consideramos como evidência confiável sobre os efeitos dos tratamentos. Em retrospecto, profissionais da saúde na maioria, se não em todos, os setores do atendimento à saúde, prejudicaram seus pacientes inadvertidamente, às vezes, em uma escala muito ampla ([clique aqui para obter exemplos](#)). Na verdade, às vezes os próprios pacientes prejudicam outros pacientes quando, com base em teorias não comprovadas e experiências pessoais limitadas, incentivam o uso de tratamentos que acabaram se comprovando prejudiciais à saúde. A questão não é se devemos culpar estas pessoas, mas se os efeitos prejudiciais dos tratamentos testados de forma inadequada podem ser reduzidos. Em grande parte podem.

Reconhecer que os tratamentos podem, às vezes, trazer mais danos do que benefícios é um pré-requisito para a redução de danos não intencionais (Gregory 1772; [Haygarth 1800](#); Fordyce 1802; [Behring 1893](#)). Temos que estar, portanto, mais preparados para admitir as incertezas sobre os efeitos de tratamentos e testar estes tratamentos para reduzir as incertezas de forma suficiente. Estes testes são os experimentos controlados.

Por que as teorias sobre os efeitos de tratamentos devem ser testadas na prática

É comum as pessoas serem prejudicadas porque os tratamentos são baseados apenas em teorias sobre como a doença deveria ser tratada, sem testar como as teorias funcionam na prática. Por exemplo, por vários séculos as pessoas acreditaram na teoria de que as doenças seriam causadas por 'desequilíbrios humorais'; os pacientes eram submetidos à sangria e à purgação, sendo forçados a vomitar e a tomar inalação, na crença de que isso eliminaria os supostos desequilíbrios, até que, já no início do século 17, um único médico flamengo teve a coragem de desafiar as autoridades médicas daquele tempo a avaliar a veracidade das suas teorias, propondo um experimento controlado dos resultados dos desagradáveis tratamentos aos quais os pacientes eram submetidos ([Van Helmont 1662](#)).



No começo do século 19, os cirurgiões das forças armadas britânicas tinham começado a mostrar os efeitos nocivos da sangria no tratamento de "febres" (Robertson 1804; [Hamilton 1816](#)).

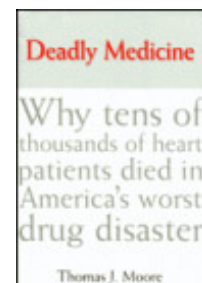
Algumas décadas mais tarde, a prática também foi contestada por um médico parisiense ([Louis 1835](#)). Já no início do século XX, os praticantes ortodoxos em Boston, EUA, que não usavam o método de sangria para tratar a pneumonia ainda eram considerados negligentes (Silverman 1980). Na verdade, Sir William Osler, uma das autoridades médicas mais influentes no mundo, geralmente cauteloso a respeito de tratamentos não comprovados, informou seus leitores de que: "ao longo das últimas décadas certamente fizemos muito poucas sangrias. A Pneumonia é uma das doenças em que uma flebotomia [sangria] feita na hora certa pode salvar uma vida. Para ser eficaz, deve ser feita logo. Em um homem forte e saudável que apresenta febre alta e pulso acelerado, a retirada de 567 a 850 gramas de sangue é sempre benéfica" (Osler 1892).



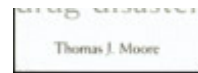
Embora a necessidade de testar a veracidade das teorias na prática ter sido reconhecida por algumas pessoas há pelo menos um milênio ([Ibn Hindi século X - XI](#)), este importante princípio ainda é, muitas vezes, ignorado. Por exemplo, baseado em uma teoria não testada, Benjamim Spock, influente especialista em pediatria norte-americano, informou aos leitores de seu livro de maior sucesso 'Baby and Child Care' que a desvantagem de os bebês dormirem de costas era que, caso vomitassem, a probabilidade deles engasgarem seria maior. O Dr. Spock aconselhou, então, seus milhares de leitores a incentivar os bebês a dormirem sobre a barriga (Spock 1966). Hoje sabemos que este conselho, aparentemente racional na teoria, levou à morte súbita de dezenas de milhares de recém nascidos (Gilbert et al. 2004).



O uso de medicamentos para prevenir as anormalidades do batimento cardíaco em pessoas que tiveram ataque cardíaco fornece um outro exemplo dos riscos da aplicação de uma teoria não comprovada na prática. Devido à associação das anormalidades do batimento cardíaco ao aumento no risco de morte prematura após um ataque cardíaco, a teoria dizia que estes medicamentos reduziriam tais mortes prematuras. Porém, só porque uma teoria parece lógica não significa que ela esteja necessariamente correta. Anos depois dos medicamentos terem sido licenciados e adotados na prática, descobriu-se que eles, na verdade, aumentavam o risco de morte súbita após um ataque cardíaco. De fato, estima-se que, no auge de seu uso no final dos anos 80, elas podem



ter matado até 70.000 pessoas por ano, só nos Estados Unidos (Moore 1995) – um número muito maior do que o de americanos que morreram na Guerra do Vietnã.



Por outro lado, a confiança equivocada no pensamento teórico como uma orientação para a prática também resultou na rejeição inadequada de alguns tratamentos, porque os pesquisadores não acreditavam que estes tratamentos pudessem funcionar. As teorias baseadas nos resultados de pesquisas com animais, por exemplo, às vezes prevêem corretamente os resultados dos experimentos dos tratamentos em seres humanos, mas este nem sempre é o caso. Baseado nos resultados de experimentos em ratos, alguns pesquisadores se convenceram que não havia razão em administrar medicamentos para dissolver coágulos em pacientes que sofreram ataques cardíacos mais de seis horas antes. Se estes pacientes não tivessem participado de alguns experimentos controlados com estes medicamentos, não saberíamos que eles podem se beneficiar com o tratamento (Fibrinolytic Therapy Trialists' Collaborative Group 1994).

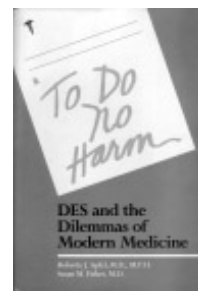


As observações na prática clínica ou em laboratório e na pesquisa com animais podem sugerir se determinados tratamentos beneficiarão ou não os pacientes, mas conforme estes e muitos outros exemplos deixam claro, é fundamental usar experimentos controlados para descobrir se, na prática, estes tratamentos fazem mais bem do que mal, ou vice-versa.

Por que os testes com tratamentos médicos devem ser experimentos controlados

A incapacidade de testar as teorias sobre tratamentos na prática não é a única causa evitável das tragédias causadas pelos tratamentos. Elas também têm ocorrido porque os testes usados para avaliar os efeitos dos tratamentos são duvidosos e equivocados. Os experimentos controlados envolvem medidas para reduzir a probabilidade de sermos iludidos pelos efeitos da [parcialidade](#) e [do acaso](#).

Por exemplo, na década de 50, a teoria e os experimentos mal controlados que produziam evidências duvidosas indicavam que ao dar um estrógeno sintético, o dietilstilbestrol (DES), para mulheres grávidas com história de aborto e natimorto aumentaria a probabilidade de um desfecho bem sucedido em gestações futuras. Embora os experimentos controlados sugerissem que o DES fosse ineficaz, a teoria e a evidências duvidosas, junto com um marketing agressivo, fizeram com que o DES fosse prescrito a milhares de mulheres grávidas durante as décadas seguintes. As conseqüências foram desastrosas: algumas das filhas das mulheres que foram receberem DES desenvolveram câncer de vagina, e outras crianças tiveram outros problemas de saúde, inclusive má-formação dos órgãos reprodutivos e infertilidade (Apfel and Fisher 1984).



Os problemas resultantes de testes de tratamentos inadequados continuam ocorrendo. Novamente, como resultado de evidências duvidosas e do marketing agressivo, milhares de mulheres foram convencidas a fazer uso da terapia de reposição hormonal (TRH), não só porque esta poderia reduzir os sintomas desagradáveis da menopausa, mas também porque reduziria suas chances de ter ataques cardíacos e derrames cerebrais. Quando estas declarações foram avaliadas em experimentos controlados os resultados mostraram que, longe de reduzir os riscos de ataques cardíacos e derrames, a TRH aumenta os riscos destas condições potencialmente fatais, assim como teria outros efeitos indesejáveis (McPherson 2004).



Estes exemplos da necessidade de experimentos controlados de tratamentos são alguns dos muitos que ilustram como os tratamentos podem fazer mais mal do que bem. São necessários maiores conhecimentos gerais sobre os experimentos controlados de tratamentos, de modo que, marcados com uma dose saudável de ceticismo, todos possamos avaliar as afirmações sobre os efeitos dos tratamentos de modo mais crítico. Desta forma, nós todos seremos mais capazes de julgar quais tratamentos são mais propensos a trazer mais benefícios do que danos.

Os princípios dos experimentos controlados de tratamentos têm evoluído há séculos, e ainda hoje continuam evoluindo.

Cite as: Editorial commentary (2007). Por que os experimentos controlados são necessários? The James Lind Library (www.jameslindlibrary.org).

Next essay: [Por que as comparações são fundamentais?](#)

Select essay:

Referências

Apfel RJ, Fisher SM (1984). To do no harm: DES and the dilemmas of modern medicine. New Haven, Ct: Yale University Press.

Behring, Boer, Kossel H (1893). Zur Behandlung diphtheriekranker Menschen mit Diphtherieheilserum. Deutsche Medicinische Wochenschrift 17: 389-393.

Fibrinolytic Therapy Trialists' Collaborative Group (1994). Indications for fibrinolytic therapy in suspected acute myocardial infarction: collaborative overview of early mortality and major morbidity results from all randomised trials of more than 1000 patients. Lancet 1994;343: 311-322.

Fordyce G (1802). A second dissertation on fever. London: J Johnson

Gilbert R, Salanti G, Harden M, See S (2005). Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of recommendations from 1940 to 2002. International Journal of Epidemiology 34:874-87.

Gregory J (1772). Lectures on the duties and qualifications of a physician. London: Strahan and Cadell.

Hamilton AL (1816). Dissertatio Medica Inauguralis De Synocho Castrensi (Inaugural medical dissertation on camp fever). Edinburgh: J Ballantyne.

Haygarth J (1800). Of the imagination, as a cause and as a cure of disorders of the body: exemplified by fictitious tractors, and epidemical convulsions. Bath: R. Crutwell.

Ibn Hindu (10th-11th century CE; 4th-5th century AH). Miftah al-tibb wa-minhaj al-tullab [The key to the science of medicine and the students' guide].

Louis PCA (1835). Recherches sur les effets de la saignée dans quelques maladies inflammatoires et sur l'action de l'émétique et des vésicatoires dans la pneumonie. Paris: Librairie de l'Académie royale de médecine.

McPherson K (2004). Where are we now with hormone replacement therapy? BMJ 328: 357-358.

Moore TJ (1995). Deadly Medicine. New York: Simon and Schuster.

Osler W (1892). Principles and Practice of Medicine. London: Appleton, p 530.

Robertson R (1804). Observations on the diseases incident to seamen, 2nd edn. Vol. 1, London: for the author.

Silverman W (1980). In: Chalmers I, McIlwaine G (eds). Perinatal Audit and Surveillance. London: Royal College of Obstetricians and Gynaecologists, 1980: 110.

Spock B (1966). Baby and Child Care. 165th printing. New York: Pocket Books, pp 163-164.

van Helmont JB (1662). Oriatrike, or physick refined: the common errors therein refuted and the whole are reformed and rectified [translated by J Chandler]. Lodowick-Loyd: London, p 526.

[Home](#)

[Conteúdo](#)

[Envie comentários](#)

Por que as comparações são fundamentais?

Será que um tratamento é melhor do que a natureza e o tempo?

Os pacientes e profissionais da saúde esperam que os tratamentos sejam eficazes. Estas expectativas otimistas podem ter um efeito muito positivo na satisfação de todos com o atendimento à saúde, conforme observou o médico britânico Richard Asher em um de seus artigos para médicos:

"Se você acreditar fervorosamente em seu tratamento, mesmo que os experimentos controlados mostrem que ele é completamente ineficaz, os seus resultados serão muito melhor, os seus pacientes se sentirão muito melhor e a sua renda será muito melhor também. Creio que isto explica o incrível sucesso de alguns membros da nossa profissão menos dotados, porém mais crédulos, e também a aversão violenta a experimentos controlados e a estatísticas que alguns médicos modernos e bem sucedidos demonstram". (Asher 1972)

É comum as pessoas se recuperarem de enfermidades sem qualquer tratamento específico: a natureza e o tempo são ótimos remédios. Conforme Oliver Wendell Holmes sugeriu no século XIX, quando havia poucos tratamentos eficazes ([Holmes 1861](#)), "Acredito firmemente que se toda a materia medica, como usada hoje, pudesse ser depositada no fundo do mar, seria tanto melhor para a humanidade, e tanto pior para os peixes."

O progresso e o desfecho da enfermidade, se deixada sem tratamento, devem obviamente ser considerados quando os tratamentos são testados: o tratamento pode melhorar ou piorar os desfechos. Os escritores por vários séculos chamaram a atenção para a necessidade de ser cético sobre as afirmações de que os efeitos de tratamentos seriam melhores do que os efeitos da natureza. Em outras palavras, " Se você deixar uma 'gripe ser curada naturalmente, provavelmente você se recuperará em uma semana; mas se você for ao médico, você se recuperará em apenas sete dias."

Efeitos do placebo

Sabendo que muitas doenças são auto-limitantes, os médicos às vezes prescrevem tratamentos inertes na esperança de que os seus pacientes obtenham benefício psicológico, o assim chamado efeito placebo. Os pacientes que acreditam que um tratamento ajudará a aliviar seus sintomas - mesmo que o tratamento, na prática, não tenha efeitos físicos - podem se sentir melhor.

Há séculos, os médicos reconhecem a importância do uso de placebos ([lista de registros relevantes](#)). Por exemplo, William Cullen relatou o uso de um placebo já no ano de 1772 ([Cullen 1772](#)), e referências a placebos aumentaram durante o século XIX ([Cummings 1805](#); [Ministry of Internal Affairs 1832](#); [Forbes 1846](#)). Austin Flint acreditava que o tratamento ortodoxo à base de medicamentos estava usurpando o crédito devido a 'natureza', por isso ele administrou um 'medicamento placebo' a treze pacientes com reumatismo, o qual consistia de um extrato altamente diluído da casca da árvore quássia. O resultado foi que "o progresso favorável dos casos foi tal que garantiu ao medicamento, no geral, a confiança total dos pacientes" ([Flint 1863](#)). No Guy's Hospital em Londres, William Withey Gull chegou a conclusões semelhantes após tratar 21 pacientes com febre reumática "em grande parte com água de hortelã" ([Sutton 1865](#)). No início do século 20, William Rivers discutiu os efeitos psicologicamente mediados dos tratamentos de forma detalhada ([Rivers 1908](#)).



A necessidade de comparações

Da mesma forma que a força curativa da natureza e o efeito placebo são reconhecidos há séculos, também sabemos da necessidade de comparações para avaliar os efeitos dos tratamentos, além dos efeitos naturais e psicologicamente mediados. Às vezes, as comparações de tratamentos são feitas nas mentes das pessoas: elas têm a impressão de que elas ou outros estão respondendo de forma diferente a um novo tratamento em comparação às reações aos tratamentos anteriores. Por exemplo, Ambroise Paré, um cirurgião militar francês, concluiu que o tratamento de ferimentos de batalha com óleo fervente (como era prática comum) era provavelmente prejudicial. Ele concluiu isso quando o fornecimento de óleo acabou e seus pacientes se recuperaram mais rápido do que o normal ([Paré 1575](#)).



Na maioria das vezes, impressões como esta precisam ser acompanhadas de perto por investigações formais, talvez começando pela análise de prontuários médicos. Tais impressões podem então levar a comparações conduzidas cuidadosamente. O perigo surge quando impressões isoladas são usadas como uma orientação para as

recomendações e decisões de tratamento.

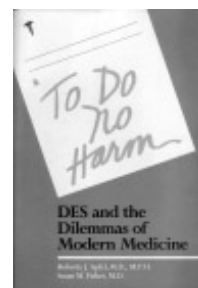
Efeitos drásticos e efeitos moderados de tratamentos

As comparações de tratamentos baseadas em impressões, ou análises relativamente restritas, somente fornecem informações confiáveis nas raras circunstâncias em que os efeitos dos tratamentos são drásticos ([clique aqui para uma lista de registros relevantes](#)). Os exemplos incluem o uso do ópio para o alívio da dor (Tibi 2005), a higiene para prevenir o tétano (tetania) em bebês recém-nascidos (Schleisner 1849), o clo(Tibi 2005)rofórmio para anestesia, a insulina para diabetes (Banting et al. 1922), a ingestão de fígado para a anemia perniciosa (Minot and Murphy 1926), medicamentos à base de sulfá para a infecção pós-parto (Colebrook and Purdie 1937), estreptomicina para a meningite tuberculosa (MRC 1948), adrenalina para as reações alérgicas potencialmente fatais (McLean-Tookey et al. 2003) e medicamentos manipulados geneticamente para algumas formas raras de leucemia (Druker et al. 2001). A maioria dos tratamentos médicos não tem efeitos tão radicais como estes, no entanto, e a menos que se tome cuidado para evitar as comparações com viés, pode-se tirar conclusões perigosamente equivocadas sobre os efeitos do tratamento.

Comparação dos tratamentos administrados hoje com os tratamentos administrados no passado

Foi em grande parte devido à confiança nas comparações com viés com experiência pregressa que os médicos e as mulheres acreditavam que o dietilstilbestrol (DES) reduziria o risco de abortos e natimortos. Nunca houve qualquer evidência a partir de experimentos controlados (imparciais) de que o DES pudesse fazer isto, e mais tarde ficou comprovado que ele causou câncer nas filhas de algumas das mulheres grávidas que receberam tal medicamento. Um tratamento cuja eficácia não foi comprovada não deve ser promovido.

A comparação de tratamentos administrados hoje com os tratamentos administrados no passado muito raramente fornece uma base sólida para um experimento controlado ([Behring et al. 1893](#); [Roux et al. 1894](#)), porque fatores relevantes além dos próprios tratamentos mudam ao longo do tempo. Por exemplo, abortos e natimortos são mais comuns nas primeiras gestações do que nas gestações posteriores. Comparar a frequência de abortos e natimortos em gestações posteriores nas quais o DES foi administrado com o desfecho das primeiras gestações na qual o medicamento não foi usado, provavelmente seja uma base equivocada para avaliar os seus efeitos. Se possível, portanto, as comparações devem envolver a aplicação de diferentes tratamentos mais ou menos ao mesmo tempo.



Comparação de tratamentos com estudos cruzados em pacientes específicos

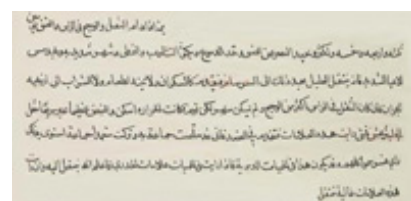
Às vezes, administrar diferentes tratamentos mais ou menos ao mesmo tempo pode implicar a administração a um paciente de diferentes tratamentos um após o outro, o assim chamado estudo cruzado ([Martini 1932](#); [clique aqui para uma lista de registros relevantes](#)). Um exemplo antigo de um estudo cruzado foi relatado em 1786 pelo Dr. Caleb Parry em Bath, na Inglaterra. Ele queria descobrir se havia alguma razão para pagar por um caro ruibarbo turco importado usado como um purgante para tratar seus pacientes, ao invés de usar o ruibarbo cultivado localmente na Inglaterra. Assim ele 'alternou' o tipo de ruibarbo administrado a cada paciente em ocasiões diferentes e depois comparou os sintomas que cada paciente experimentou ao comer cada tipo de ruibarbo ([Parry 1786](#)). (Ele não encontrou qualquer vantagem em relação ao ruibarbo caro!)



As comparações de tratamentos em pacientes específicos têm o seu lugar quando houver um retorno da condição após o fim do tratamento. Existem muitas circunstâncias às quais isto não se aplica. Por exemplo, é quase impossível comparar diferentes operações cirúrgicas desta maneira, ou tratamentos administrados em condições agravantes.

Comparação de grupos de pacientes que receberam tratamentos diferentes ao mesmo tempo

Os tratamentos são geralmente testados ao se comparar grupos de pessoas que receberam diferentes tratamentos. A comparação de dois tratamentos será parcial se pessoas relativamente saudáveis receberem um tratamento e pessoas relativamente doentes receberem outro, assim devemos comparar as experiências de grupos semelhantes de pessoas que recebem tratamentos diferentes ao longo do mesmo período de tempo. Al-Razi reconheceu isto mais de mil anos atrás, quando desejando chegar a uma conclusão sobre como tratar os pacientes com sinais de início de meningite, ele tratou um grupo de pacientes e intencionalmente impediu o tratamento de um grupo de comparação ([al-Razi 9th century](#)).



As comparações com a natureza ou com outros tratamentos são necessárias nos experimentos controlados de tratamentos. Para estas comparações serem plausíveis, elas devem [abordar incertezas legítimas](#), evitar os [vieses](#) e o [efeito do caso](#), além de [serem interpretadas cuidadosamente](#).

Cite as: Editorial commentary (2007). Por que as comparações são fundamentais? The James Lind Library (www.jameslindlibrary.org).

Next essay: [Por que as comparações devem tratar de dúvidas legítimas?](#)

Select essay:

Referências

al-Razi (10th century CE; 4th Century AH). Kitab al-Hawi fi al-tibb [The comprehensive book of medicine].

Asher R (1972). Talking sense. London: Pitman Medical.

Banting FG, Best CH, Collip JB, Campbell WR, Fletcher AA (1922). Pancreatic extracts in the treatment of diabetes mellitus. Canadian Medical Association Journal 12: 141-146.

Behring, Boer, Kossel H (1893). Zur Behandlung diphtheriekranker Menschen mit Diphtherieheilserum. Deutsche Medicinische Wochenschrift 17: 389-393.

Colebrook L, Purdie AW (1937). Treatment of 106 cases of puerperal fever by sulphanilamide. Lancet 2: 1237-1242 & 1291-1294.

Cullen W (1772). Clinical lectures. Edinburgh, February-April, 218-9.

Cummngs R (1805). Medical and Physical Journal, page 6.

Druker BJ, Talpaz M, Resta DJ, Peng B, Buchdunger E, Ford JM, Lydon NB, Kantarjian H, Capdeville R, Ohno-Jones S, Sawyers CL (2001). Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. New England Journal of Medicine 344: 1031-1037.

Flint A (1863). A contribution toward the natural history of articular rheumatism; consisting of a report of thirteen cases treated solely with palliative measures. American Journal of the Medical Sciences 46: 17-36.

Forbes J (1846). Homeopathy, allopathy and 'young physic.' British and Foreign Medical Review 21: 225-265.

Holmes OW (1861). Currents and countercurrents in medical science. In: Works, 1861 Vol ix, p 185.

Martini P (1932). Methodenlehre der Therapeutischen Untersuchung. Berlin: Springer.

McLean-Tooke APC, Bethune CA, Fay AC, Spickett GP (2003). Adrenaline in the treatment of anaphylaxis: what is the evidence? BMJ 327: 1332-1335.

Medical Research Council (1948). Streptomycin treatment of tuberculous meningitis. Lancet 1: 582-596 .

Ministry of Internal Affairs (1823). [Conclusion of the Medical Council regarding homeopathic treatment]. Zhurnal Ministerstva Vnutrennih del, 3: 49-63.

Minot GR, Murphy WP (1926). Treatment of pernicious anaemia by a special diet. JAMA 87: 470-476.

Paré A (1575). Les oeuvres de M. Ambroise Paré conseiller, et premier chirugien du Roy avec les figures & portraits tant de l'Anatomie que des instruments de Chirugie, & de plusieurs Monstres. Paris: Gabriel Buon.

Parry CH (1786). Experiments relative to the medical effects of Turkey Rhubarb, and of the English Rhubarbs, No. I and No. II made on patients of the Pauper Charity. Letters and Papers of the Bath Society III: 431-453.

Rivers WHR (1908). The influence of alcohol and other drugs on fatigue. London: Edward Arnold.

Roux E, Martin L, Chaillou A (1894). Trois cent cas de diphthérie traité par le serum antidiphthérique. Annales de l'Institut Pasteur 8: 640-661.

Schleisner PA (1849). *Island fra et lægevidenskabeligt Synspunkt*. København: Boghandler Iversen.

Sutton HG (1865). Cases of rheumatic fever, treated for the most part by mint water. Collected from the clinical books of Dr Gull, with some remarks on the natural history of that disease. *Guy's Hospital Report* 11:392-428.

Tibi S (2005). *The medicinal use of opium in ninth-century Baghdad*. Leiden: Brill.

[Home](#)

[Conteúdo](#)

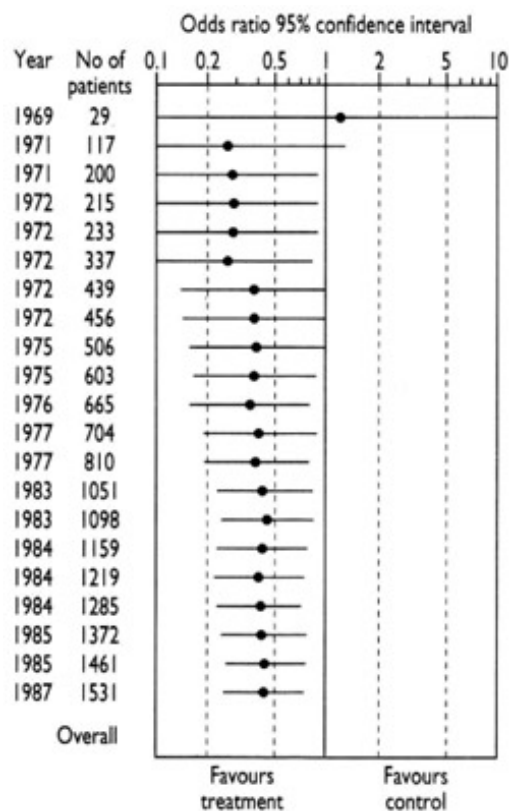
[Envie comentários](#)

[Home](#)[Conteúdo](#)jameslindlibrary.org

Por que as comparações devem tratar de dúvidas legítimas?

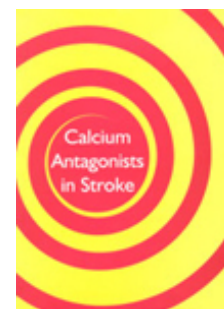
Uma grande quantidade de pesquisas são feitas mesmo quando não existem incertezas legítimas. Os pesquisadores que falham na condução de avaliações sistemáticas de testes passados antes de iniciar novos estudos às vezes não reconhecem (ou preferem ignorar o fato) que as incertezas sobre os efeitos do tratamento já foram tratadas de forma convincente. Isto significa que as pessoas que participam da pesquisa, às vezes, não têm o tratamento que poderia ajudá-las, ou recebem um tratamento que pode prejudicá-las.

O diagrama que acompanha este parágrafo e o seguinte mostra o acúmulo de evidência a partir de experimentos controlados realizados com o intuito de avaliar se os antibióticos (comparados com placebos inativos) reduzem o risco de morte pós-operatória em pessoas submetidas à cirurgia de intestino (Lau et al. 1995). O primeiro experimento controlado foi relatado em 1969. Os resultados deste pequeno estudo deixaram incertezas sobre a eficiência dos antibióticos, a linha horizontal representando os resultados transpõe a linha vertical que separa os efeitos favoráveis e desfavoráveis dos antibióticos. De forma bem oportuna, essa incerteza foi tratada em outros experimentos no início dos anos 70.



Conforme as evidências se acumulavam, entretanto, ficou evidente na metade dos anos 70 que os antibióticos reduzem o risco de morte após a cirurgia (a linha horizontal cai claramente ao lado da linha vertical, favorecendo o tratamento). Mesmo assim os pesquisadores continuaram a realizar estudos até o final dos anos 80. Metade dos pacientes que receberam placebos nestes novos estudos não puderam receber uma forma de tratamento que revelou reduzir o risco de eles morrerem após a cirurgia. Como isto aconteceu? Provavelmente porque os pesquisadores continuaram com as pesquisas sem revisar as evidências existentes sistematicamente. Este comportamento permanece bastante comum na comunidade de pesquisa, em grande parte porque alguns dos incentivos no mundo da pesquisa – comercial e acadêmica – não coloca os interesses dos pacientes em primeiro lugar (Chalmers 2000).

Os pacientes e participantes da pesquisa também podem ser prejudicados porque os pesquisadores não fizeram uma revisão sistemática das evidências relevantes na pesquisa com animais antes de começarem a testar os tratamentos em seres humanos. Uma equipe holandesa revisou a experiência de mais de 7.000 pacientes que participaram dos testes de um novo medicamento para bloqueio do cálcio administrado em pessoas que sofreram derrame cerebral. Eles não encontraram nenhuma evidência que sustentasse o crescente uso deste medicamento na prática (Horn and Limburg 2001). Isso fez com que eles pensassem sobre a qualidade e os achados da pesquisa com animais que levou à pesquisa com pacientes. Suas revisões dos estudos com animais revelaram que nunca se sugeriu que o medicamento seria benéfico aos seres humanos (Horn et al. 2001).



A razão mais comum do porquê a pesquisa não trata de incertezas legítimas é o fato de os pesquisadores simplesmente não serem suficientemente disciplinados para revisar sistematicamente as evidências existentes relevantes antes de iniciar novos estudos. Porém, às vezes, há razões mais sinistras. Os pesquisadores podem estar cientes das evidências existentes, mas eles querem planejar estudos para assegurar que suas próprias pesquisas produzirão resultados favoráveis para tratamentos específicos. Geralmente, mas em sempre, isto ocorre

por razões comerciais (Djulbegovic et al. 2000; Sackett and Oxman 2003). Esses estudos são deliberadamente concebidos para serem experimentos não controlados de tratamentos. Isto por ser feito pela ocultação de um tratamento de comparação reconhecido por ajudar os pacientes (como no exemplo dado acima), ou pela administração de tratamentos de comparação em pequenas doses de forma inapropriada (de modo que não funcionem tão bem), ou em doses muito altas (de modo que tenham mais efeitos colaterais indesejados) ([veja o comentário de Mann e Djulbegovic](#)). E, também pode resultar do acompanhamento dos pacientes por um período de tempo muito curto (deixando passar despercebidos os efeitos tardios do tratamento) e do uso de medidas de desfecho ("substitutas") que têm pouca ou nenhuma correlação com os desfechos que importam para os pacientes.

Os leitores deste artigo podem ficar surpresos ao saber que os comitês de ética na pesquisa, estabelecidos durante as últimas décadas para garantir que a pesquisa seja ética, têm feito muito pouco para influenciar esta negligência nas pesquisas. A maioria destes comitês tem desapontado as pessoas que eles deveriam proteger, porque não exigem dos pesquisadores e patrocinadores procurando a aprovação de novos experimentos uma revisão sistemática das evidências existentes (Savulescu et al. 1996; Chalmers 2002). A falha dos comitês de ética em proteger os pacientes e o público de maneira eficiente, enfatiza a importância de melhorar o conhecimento geral sobre as características dos experimentos controlados de tratamentos médicos.

Cite as: Editorial commentary (2007). Por que as comparações devem tratar de dúvidas legítimas? The James Lind Library (www.jameslindlibrary.org).

Next essay: [Precaução nas comparações com vícios](#)

Select essay:

Referências

Chalmers I. Current Controlled Trials: an opportunity to help improve the quality of clinical research. *Current Controlled Trials in Cardiovascular Medicine* 2000;1:3-8.

Disponível no site: <http://cvm.controlled-trials.com/content/1/1/3>

Chalmers I (2002). Lessons for research ethics committees. *Lancet* 359:174.

Djulbegovic B, Lacey M, Cantor A, Fields KK, Bennett CL, Adams JR, Kuderer NM, Lyman GH (2000). The uncertainty principle and industry-sponsored research. *Lancet* 356:635-638 .

Horn J, Limburg M (2001). Calcium antagonists for acute ischemic stroke (Cochrane Review). In: *The Cochrane Library*, Issue 3, Oxford: Update Software.

Horn J, de Haan RJ, Vermeulen M, Luiten PGM, Limburg M (2001). Nimodipine in animal model experiments of focal cerebral ischaemia: a systematic review. *Stroke* 32:2433-38.

Lau J, Schmid CH, Chalmers TC (1995). Cumulative meta-analysis of clinical trials builds evidence for exemplary clinical practice. *Journal of Clinical Epidemiology* 48:45-57.

Mann H, Djulbegovic B. Why comparisons must address genuine uncertainties. James Lind Library (www.jameslindlibrary.org).

Sackett DL, Oxman AD (2003). HARLOT plc: an amalgamation of the world's two oldest professions. *BMJ* 2003;327:1442-1445.

Savulescu J, Chalmers I, Blunt J (1996). Are research ethics committees behaving unethically? Some suggestions for improving performance and accountability. *BMJ* 313:1390-1393.

[Home](#)

[Conteúdo](#)

[Envie comentários](#)

[Home](#)[Conteúdo](#)[jameslindlibrary.org](http://www.jameslindlibrary.org)

Precaução nas comparações com vícios

Algumas vezes os tratamentos têm efeitos drásticos ([clique aqui uma lista de registros relevantes](#)) Estes podem ser indesejados e específicos, por exemplo, quando uma pessoa tem uma reação alérgica a um medicamento antibiótico. Os tratamentos também podem ter efeitos benéficos impressionantes, como a adrenalina nas reações alérgicas potencialmente fatais (McLean-Tooke et al. 2003). Porém, efeitos tão impressionantes são raros. Geralmente, os efeitos dos tratamentos são mais modestos, mesmo assim devemos saber sobre eles, por exemplo, o uso da aspirina para reduzir o risco de ataque cardíaco ([Elwood 2004](#)).

Por exemplo, a aspirina não previne todas as mortes prematuras após um ataque cardíaco, mas reduz a probabilidade de morte em cerca de vinte por cento, o que é importante em uma condição tão comum. Se estes efeitos moderados, porém importantes da maioria dos tratamentos forem detectados com segurança, devemos tomar cuidado para assegurar que comparações com viés não nos levem a acreditar que os tratamentos são eficazes quando na verdade são ineficazes ou prejudiciais, ou ineficazes quando eles podem na verdade ser eficazes.

Os vieses nos experimentos de tratamentos são aquelas influências e fatores que podem levar a conclusões sobre os efeitos dos tratamentos que diferem totalmente da verdade. Embora muitos tipos de vieses possam distorcer os resultados de pesquisas em saúde (Sackett 1979), na James Lind Library nos concentramos naqueles vieses que devem ser minimizados nos experimentos controlados de tratamentos. Estes são:

- [vieses devido a diferenças nas pessoas comparadas](#);
- [vieses devido a diferenças na maneira como os desfechos dos tratamentos são avaliados](#);
- [viés de informação da evidência disponível](#); e
- [seleção com viés a partir da evidência disponível](#).

Ignorar estes vieses (ou às vezes, inescrupulosamente, se aproveitar deles), pode levar as pessoas a acreditar que um novo tratamento é melhor do que um tratamento existente, quando na verdade não é. Isto pode ocorrer em decorrência de basear as conclusões em:

- estudos que comparam o progresso de pessoas relativamente saudáveis que recebem um novo tratamento ao progresso de pessoas relativamente doentes que recebem um tratamento padrão ([viés da distribuição](#)).
- estudos em que as avaliações dos desfechos do tratamento provavelmente estejam com viés em favor de um novo tratamento, por exemplo, comparando as opiniões de pessoas que sabem que receberam um novo e caro tratamento com as opiniões daquelas que podem ter se decepcionado porque continuaram a usar um tratamento padrão desestimulante ([viés do observador ou da medida](#)).
- somente estudos que mostram um novo tratamento em um aspecto favorável, e não aqueles que sugerem que o tratamento possa ser prejudicial, que muitas vezes não são relatados ([viés da informação](#)).
- seleção e interpretação da evidência disponível com viés para justificar um ponto de vista em particular ([viés da revisão](#)).

Geralmente, os experimentos não controlados que resultam destes vieses não são reconhecidos pelo o que são. Entretanto, é comum pessoas com interesses escusos explorarem estes vieses de modo que os tratamentos sejam apresentados como se eles fossem melhor do que realmente são (Sackett and Oxman 2003).

Sendo os vieses involuntários ou propositais, as conseqüência são as mesmas: a não ser que os experimentos de tratamentos sejam controlados, alguns tratamentos ineficazes ou prejudiciais parecerão eficazes, enquanto que alguns tratamentos eficazes parecerão ineficazes ou prejudiciais.

Cite as: Editorial commentary (2007). Precaução nas comparações com vícios. The James Lind Library (www.jameslindlibrary.org).

Next essay: [Diferenças entre as pessoas comparadas](#)

Select essay:

Referências

Elwood P (2004). The first randomised trial of aspirin for heart attack and the advent of systematic overviews of trials. The James Lind Library (www.jameslindlibrary.org).

McLean-Tooke APC, Bethune CA, Fay AC, Spickett GP (2003). Adrenaline in the treatment of anaphylaxis: what is the evidence? *BMJ* 327:1332-1335.

Sackett DL (1979). Bias in analytic research. *Journal of Chronic Diseases* 32:51-63.

Sackett DL, Oxman AD (2003). HARLOT plc: an amalgamation of the world's two oldest professions. *BMJ* 2003;327:1442-1445.

[Home](#)

[Conteúdo](#)

[Envie comentários](#)

Home

Conteúdo

jameslindlibrary.org

Precaução contra comparações com viés:

Diferenças entre as pessoas comparadas

Comparação de diferentes tratamentos administrados a grupos de pessoas

As comparações de tratamentos geralmente envolvem a comparação das experiências dos grupos de pessoas que receberam tratamentos diferentes. Se estas comparações forem feitas adequadamente, a composição dos grupos deve ser igual, de modo que semelhante seja comparado a semelhante. Se aqueles que recebem um tratamento são mais propensos a se saírem bem (ou mal) do que aqueles que recebem um tratamento alternativo, este viés de distribuição torna impossível a certeza de que os desfechos irão refletir os efeitos diferenciais dos tratamentos, ou os efeitos da natureza e da passagem do tempo.

No século XVIII, o cirurgião inglês William Cheselden sabia do problema dos "grupos desiguais" quando os cirurgiões comparavam suas respectivas taxas de mortalidade após operações para remover cálculos renais. Cheselden afirmava que era importante considerar as idades das pessoas tratadas por cirurgiões diferentes. Ele chamou a atenção para o fato de que as taxas de mortalidade variavam com as idades dos pacientes ([Cheselden 1740](#)), pacientes com idade mais avançada tinham mais probabilidade de morrer do que os pacientes mais jovens. Isto significava que, se alguém desejasse comparar a frequência das mortes em grupos de pacientes que se submeteram a diferentes tipos de operações, deveria considerar a diferença de idade dos pacientes nos grupos de comparação.



A comparação das experiências e dos desfechos de pacientes que por acaso receberam tratamentos diferentes no passado é usada até hoje como uma forma de tentar avaliar os efeitos dos tratamentos. O desafio é saber se os grupos de comparação eram suficientemente parecidos antes de receberem o tratamento. Isto é demonstrado pelas tentativas de avaliar os efeitos da terapia de reposição hormonal (TRH), ao comparar as doenças das mulheres que usaram a TRH com as doenças de outras mulheres que não usaram a TRH. Conforme as análises subseqüentes dos experimentos controlados da TRH revelaram, tentar avaliar os efeitos dos tratamentos, em retrospecto, desta maneira, pode, muitas vezes, levar a resultados perigosamente equivocados (McPherson 2004).

É quase impossível ter total certeza de que os grupos de comparação selecionados entre as pessoas que receberam um tratamento no passado, sejam compatíveis, em todos os aspectos que importam, com as pessoas que mais recentemente receberam um tratamento alternativo. Isto se aplica mesmo que algumas informações sobre os pacientes que receberam tratamentos diferentes estejam disponíveis (tais como a idade, ou o histórico de doenças). Outras informações que podem ser de grande importância (como a probabilidade de recuperação espontânea) podem simplesmente não estar disponíveis.

Um abordagem melhor é planejar as comparações dos tratamentos antes de iniciar um tratamento. Por exemplo, antes de iniciar sua comparação entre seis tratamentos para escorbuto a bordo do navio *HMS, em Salisbury*, em 1747, James Lind teve o cuidado de selecionar pacientes que estivessem em um estágio semelhante desta doença muitas vezes fatal. Ele também assegurou-se de que tinham a mesma dieta básica e estavam acomodados em condições semelhantes. Estes eram fatores, além do tratamento, que poderiam ter influenciado suas chances de recuperação ([Lind 1753](#)). Deve-se empregar esforços parecidos para tentar garantir que os grupos de comparação dos tratamentos sejam compostos de pessoas semelhantes.



Montagem imparcial dos grupos de comparação dos tratamentos usando alternância ou randomização

Apesar de James Lind ter tomado o cuidado de assegurar que os marinheiros nos seis grupos de comparação fossem semelhantes, ele não descreve como decidiu quais marinheiros receberiam quais dos seis tratamentos. Existe apenas uma maneira de assegurar que os grupos de comparação dos tratamentos sejam definidos de forma que eles sejam parecidos em todas as maneiras que interessam, conhecidas e desconhecidas. E esta é através do uso de alguma forma de processo casual para formar os grupos de comparação de tratamentos, evitando assim a seleção com viés para diferentes tratamentos antes de iniciar o tratamento.

Cem anos após Lind, um médico militar, Graham Balfour, demonstrou como isto poderia ser feito em um experimento para observar se a beladona prevenia a escarlatina em crianças. No orfanato militar do qual ele era responsável, ele usou a alternância, "para evitar a atribuição de seleção", para decidir quais garotos receberiam e quais não receberiam a beladona ([Balfour](#)





1854). A alternância é um dos vários métodos imparciais usados para formar grupos semelhantes de comparação de tratamentos antes de administrar os tratamentos que estão sendo comparados. Durante a primeira metade do século XX, existem muitos exemplos de grupos de comparação de tratamentos sendo formados utilizando a alternância ou a rotação (por exemplo [Hamilton 1816](#); [MRC 1944](#)), ou por sorteio ([Colebrook 1929](#)), por exemplo, usando dados ([Doull et al. 1931](#)), contas coloridas ([Theobald 1937](#)) ou números de amostragem aleatória ([Bell 1941](#); [MRC 1948](#); [MRC 1950](#); [MRC 1951](#)). Esta 'alocação randômica' é a única, mas crucialmente importante, característica da categoria de experimentos controlados referidos como 'randomizados'. Uma alocação randômica (diferente de acidental) significa que as chances de algo acontecer são conhecidas, mas os resultados não podem ser antecipados em qualquer ocasião específica. Assim, por exemplo, se uma moeda é usada para randomizar, a probabilidade de se ter cara é de 50%, mas é impossível saber qual será o resultado de um sorteio específico.

Conforme demonstrado no artigo disponível ao [clique aqui](#), tirar a sorte é uma maneira tradicional de tomar decisões imparciais. Estes métodos ajudam a garantir que os grupos de comparação não sejam compostos de diferentes tipos de pessoas. Fatores de importância conhecidos e medidos, como a idade, podem ser verificados. Entretanto, aqueles fatores não podem ser medidos e que podem influenciar na recuperação da doença, como a dieta, o trabalho e a ansiedade, normalmente tendem a se compensar. Se desejar observar como a alocação randômica gera grupos semelhantes de pessoas ([clique aqui para uma demonstração](#)).

Conforme as experiências de uso da alternância e da alocação randômica para a montagem imparcial de grupos de pacientes na comparação de diferentes tratamentos ficava mais comum, tornou-se claro que a adesão estrita aos programas de alocação era uma exigência para evitar a criação de grupos de comparação de tratamentos com viés ([MRC 1934](#)). O risco da alocação com viés pode ser abolido se os programas de alocação dos tratamentos forem ocultados daqueles que tomam as decisões sobre a participação nas comparações dos tratamentos, em resumo, para prevenir que eles trapaceiem e, por conseguinte, induzam as comparações ([MRC 1944](#); [MRC 1948](#); [MRC 1950](#); [MRC 1951](#)).



Precaução contra perdas com viés a partir dos grupos de comparação de tratamentos

Após tomar todas as medidas para garantir que os grupos de comparação dos tratamentos sejam formados de forma a garantir que semelhantes sejam comparados a semelhantes, é importante evitar que o viés seja introduzido como um resultado da retirada seletiva de pacientes dos grupos de comparação. O quanto possível, a similaridade dos grupos deve ser mantida garantindo que todas as pessoas alocadas nos grupos de comparação de tratamentos sejam acompanhadas e incluídas na análise principal dos resultados dos testes – também conhecida como a análise da 'intenção de tratar' ([Bell 1941](#)).

Uma falha neste ponto pode resultar em experimentos não controlados de tratamentos. Tome, por exemplo, duas maneiras muito diferentes de tratar pessoas que têm ataques de vertigem em razão de vasos sanguíneos que abastecem seus cérebros parcialmente bloqueados. O tratamento para esta condição pode ser importante porque estas pessoas que sofrem de vertigem correm um risco maior de sofrer um derrame, que pode deixá-las incapacitadas, ou até matá-las. Um dos tratamentos para ataques de vertigem envolve tomar aspirina para impedir que o bloqueio fique pior; o outro envolve um procedimento cirúrgico para tentar remover o bloqueio no vaso sanguíneo.

Uma comparação controlada destas duas abordagens para tratar os ataques de vertigem implica na criação de dois grupos de pessoas usando um método de alocação imparcial (como a randomização). A comparação, portanto, iniciaria ao comparar dois grupos de pacientes semelhantes, e partir para a comparação das respectivas frequências dos derrames subsequentes. Mas se a frequência de derrames no grupo tratado cirurgicamente fosse registrada somente entre os pacientes que sobreviveram aos efeitos imediatos da operação, não perceberíamos o importante fato de que a própria operação pode causar derrame e morte. Isto resultaria em uma comparação distorcida dos dois tratamentos, gerando uma imagem falsamente otimista e com viés dos efeitos da operação. Assim, semelhante não estaria sendo comparado com semelhante.

A principal comparação nos estudos randomizados deve ser baseada, o mais possível, em todas as pessoas designadas para receber cada um dos tratamentos comparados, sem exceções, e nos grupos aos quais foram originalmente designadas. Se este princípio não for observado, as pessoas podem receber informações com viés sobre os efeitos gerais dos tratamentos.

Cite as: Editorial commentary (2007). Diferenças entre as pessoas comparadas. The James Lind Library (www.jameslindlibrary.org).

Show JLL records: illustrating [control of allocation bias](#) and [crossover test](#)

Next essay: [Diferenças na maneira como os resultados dos tratamentos são avaliados](#)

Select essay:

Referências

Balfour TG (1854). Quoted in West C. Lectures on the Diseases of Infancy and Childhood. London, Longman, Brown, Green and Longmans, p 600.

Bell JA (1941). Pertussis prophylaxis with two doses of alum-precipitated vaccine. Public Health Reports 56:1535-1546.

Cheselden W (1740). The anatomy of the human body. 5th edition. London: William Bowyer.

Colebrook D (1929). Irradiation and health. Medical Research Council Special Report Series No.131.

Doull JA, Hardy M, Clark JH, Herman NB (1931). The effect of irradiation with ultra-violet light on the frequency of attacks of upper respiratory disease (common colds). American Journal of Hygiene 13:460-77.

Hamilton AL (1816). Dissertatio Medica Inauguralis De Synocho Castrensi (Inaugural medical dissertation on camp fever). Edinburgh: J Ballantyne.

Lind J (1753). A treatise of the scurvy. In three parts. Containing an inquiry into the nature, causes and cure, of that disease. Together with a critical and chronological view of what has been published on the subject. Edinburgh: Printed by Sands, Murray and Cochran for A Kincaid and A Donaldson.

McPherson K (2004). Where are we now with hormone replacement therapy? BMJ 328:357-358.

Medical Research Council Therapeutic Trials Committee (1934). The serum treatment of lobar pneumonia. BMJ 1:241-245.

Medical Research Council (1944). Clinical trial of patulin in the common cold. Lancet 2:373-5 .

Medical Research Council (1948). Streptomycin treatment of pulmonary tuberculosis: a Medical Research Council investigation. BMJ 2:769-782.

Medical Research Council (1950). Clinical trials of antihistaminic drugs in the prevention and treatment of the common cold. BMJ 2:425-431.

Medical Research Council (1951). The prevention of whooping-cough by vaccination. BMJ 1:1463-1471.

Parry CH (1786). Experiments relative to the medical effects of Turkey Rhubarb, and of the English Rhubarbs, No. I and No. II made on patients of the Pauper Charity. Letters and Papers of the Bath Society III:407-422.

Silverman WA, Chalmers I. Casting and drawing lots. The James Lind Library (www.jameslindlibrary.org).

Theobald GW (1937). Effect of calcium and vitamin A and D on incidence of pregnancy toxæmia. Lancet 2:1397-1399 .

[Home](#)

[Conteúdo](#)

[Envie comentários](#)

[Home](#)[Conteúdo](#)jameslindlibrary.org

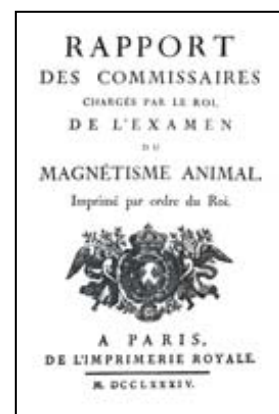
Precaução contra comparações com viés:

Diferenças na maneira como os resultados dos tratamentos são avaliados

O uso de mascaramento para reduzir o viés na avaliação dos desfechos de tratamentos

Em alguns desfechos usados para avaliar um tratamento, sobrevivência, por exemplo, uma avaliação com viés é muito pouco provável porque há pouco espaço para discussão. Este foi o caso em alguns experimentos de procedimentos cirúrgicos do século XVIII, quando a sobrevivência era a principal medida de sucesso ou fracasso do tratamento ([Faure 1759](#)). A avaliação da maioria dos desfechos, entretanto, sempre, ou quase sempre, envolve subjetividade (como no caso dos sintomas dos pacientes). Os vieses que levam a essas impressões equivocadas são chamados de vieses do observador. Eles podem causar problemas quando as pessoas acreditam que já 'sabem' o efeito de um tratamento, ou quando possam ter motivos específicos para preferir um dos tratamentos comparados. Quando não se tomam medidas para reduzir as avaliações dos desfechos com viés nas comparações de tratamentos, os efeitos do tratamento tendem a ter uma avaliação supervalorizada (Schulz et al. 1995). Quanto maior o elemento de subjetividade na avaliação dos desfechos, maior a necessidade de reduzir estes vieses do observador para garantir experimentos controlados dos tratamentos.

Nestas circunstâncias comuns, o 'mascaramento' dos pacientes e dos médicos é um elemento desejável dos experimentos controlados. A primeira avaliação cega (mascarada) de um tratamento parece ter sido realizada por uma comissão de inquérito indicada por Luís XVI em 1784 para investigar as afirmações de Anton Mesmer sobre os efeitos do 'magnetismo animal' ([Commission Royale 1784](#)). A comissão avaliou se os supostos efeitos deste novo método de cura se deviam a qualquer força "real", ou a 'ilusões da mente'. Foi dito às pessoas, que tinham seus olhos vendados, que elas estavam recebendo ou não magnetismo quando de fato, às vezes, o contrário estava acontecendo. As pessoas que estavam sendo estudadas sentiram os efeitos do 'magnetismo animal' somente nos casos em que lhes foi dito que elas estavam recebendo o tratamento, do contrário não (Kaptchuk 1998; Schulz et al. 2002).



O uso de placebos para atingir o mascaramento

Alguns anos após os testes dos efeitos do magnetismo animal, John Haygarth realizou um experimento usando um dispositivo simulado (um placebo) para atingir o mascaramento ([Haygarth 1800](#)). A charge que acompanha este parágrafo mostra um médico tratando um paciente saudável com um dispositivo patenteado e comercializado por Elisha Perkins. Perkins afirmava que seus 'tractors' – pequenas varetas de metal - curavam uma variedade de doenças através da 'força eletrofísica'. Em um folheto intitulado 'Of the imagination as a cause and as a cure of disorders of the body: exemplified by fictitious tractors' (Da imaginação como uma causa e uma cura das desordens do corpo: exemplificado pelos "tractors" fictícios), John Haygarth relatou como ele submeteu as afirmações de Perkins a um experimento controlado. Em uma série de pacientes que não sabiam dos detalhes de sua avaliação, ele usou um estudo cruzado para comparar os tractors metálicos patenteados (os quais deviam funcionar através da 'força eletrofísica') com 'tractors' de madeira que idênticos aos originais ('tractors placebo'). Ele foi incapaz de detectar qualquer benefício dos tractors metálicos ([Haygarth 1800](#)).



O experimento controlado de John Haygarth dos tractors de Perkins é um exemplo precoce do uso de placebos para conseguir um mascaramento para reduzir os vieses na avaliação do desfecho dos tratamentos. Os placebos se tornaram uma ferramenta de pesquisa nos debates sobre homeopatia, outra importante forma de cura não convencional do século XIX. Os homeopatas em muitos casos usaram a avaliação cega e controles com placebo em suas "provas", que testavam os efeitos dos seus medicamentos em voluntários saudáveis ([Löhner 1835](#); Kaptchuk 1998). Um dos experimentos controlados por placebos mais sofisticados ocorreu na Milwaukee Academy of Medicine (Academia de Medicina de Milwaukee) em 1879-1880. Este estudo era do tipo 'duplo-cego': tanto os pacientes quanto os experimentadores ficaram alheios em relação a se o tratamento era um medicamento homeopático verdadeiro ou uma pílula de açúcar ([Storke et al. 1880](#)).

Somente muito mais tarde que uma atitude mais cética na comunidade médica levou a um reconhecimento de que havia uma necessidade de adotar avaliações cegas e



placebos para avaliar a validade das suas próprias asserções. Inspirados principalmente por farmacologistas, pesquisadores alemães foram adotando gradualmente as avaliações cegas. Por exemplo, em 1918, Adolf Bingel relatou que tentou ser "o mais objetivo possível" ao comparar dois tratamentos diferentes de difteria ([Bingel 1918](#)). Ele avaliou se ele ou seus colegas poderiam supor quais pacientes tinham recebido qual tratamento: "Não confiei apenas no meu próprio julgamento, mas busquei as opiniões dos médicos assistentes do departamento de difteria, sem informá-los sobre a natureza do soro que estava sendo testado. O julgamento deles foi, portanto, livre de viés. Fico feliz de ver minhas observações verificadas independentemente, e sem dúvidas recomendo este método 'cego' para o propósito." ([Bingel 1918](#)). Na verdade, nenhuma diferença foi detectada entre os dois tratamentos. Uma forte tradição de avaliação cega se desenvolveu na Alemanha, e foi sistematizada pelo farmacologista clínico Paul Martini ([Martini 1932](#)).



A avaliação cega em países de língua inglesa modernos começou, a princípio, quando os farmacologistas influenciados pela tradição alemã, assim como por um movimento nativo chamado de 'quackbuster' (caçadores de fraudes) que usava avaliações cegas (Kaptchuk 1998). Na década de 30, eles conquistaram a liderança no uso de controles com placebo em experimentos clínicos. Por exemplo, dois dos primeiros experimentos controlados do Conselho de Pesquisa Médica do Reino Unido (UK Medical Research Council) eram de tratamentos para resfriado. Teria sido muito difícil interpretar os resultados se o método 'duplo cego' não tivesse sido usado para impedir que os pacientes e os médicos soubessem quais pacientes tinham recebido os novos medicamentos e quais tinham recebido placebos ([MRC 1944](#); [MRC 1950](#)). A defesa tenaz de Harry Gold da importância da avaliação cega parece ter tido uma influência fundamental nos Estados Unidos ([Conference on Therapy 1954](#)).

Cegar os observadores quando é impossível cegar pacientes e médicos

Às vezes é simplesmente impossível cegar os pacientes e os médicos sobre a identidade dos tratamentos comparados, por exemplo, quando tratamentos cirúrgicos são comparados a tratamentos com medicamentos ou com nenhum tratamento. Mesmo nestas circunstâncias, todavia, podem ser tomadas medidas para reduzir a avaliação com viés dos desfechos dos tratamentos. Observadores independentes podem ficar alheios a quais tratamentos foram recebidos por quais pacientes. Por exemplo, no início da década de 40 um experimentou comparou pacientes com tuberculose pulmonar recebendo o então tratamento padrão - repouso - a outros pacientes que receberam, além disso, injeções do antibiótico estreptomicina. Os pesquisadores acharam que seria antiético injetar placebos inativos nos pacientes alocados para repouso simplesmente para atingir o 'mascaramento' dos pacientes e médicos que os tratavam ([MRC 1948](#)), mas eles tomaram precauções alternativas para reduzir a avaliação com viés dos desfechos. Embora houvesse pouco perigo de avaliação com viés do resultado principal (sobrevivência), a subjetividade poderia ter induzido a avaliação dos raios-X do tórax. Assim sendo, os raios-X foram analisados por médicos que não sabiam se eles estavam avaliando o resultado em um paciente que tinha sido tratado com estreptomicina ou somente com repouso.

Junto com a randomização, a avaliação cega, quando possível usando placebos, agora se tornou um dos componentes metodológicos cruciais dos experimentos controlados de tratamentos.

Cite as: Editorial commentary (2007). Diferenças na maneira como os resultados dos tratamentos são avaliados. The James Lind Library (www.jameslindlibrary.org).

Show JLL records: illustrating [control of observer bias](#)

Next essay: [Interpretação de comparações sem vícios](#)

Select essay:

Referências

Bingel A (1918). Über Behandlung der Diphtherie mit gewöhnlichem Pferdeserum. Deutsches Archiv für Klinische Medizin 125:284-332.

Commission Royale (1784). Rapport des commissaires chargés par le roi du magnetisme animal. Paris: Imprimerie royale.

Conference on Therapy (1954). How to evaluate a new drug. American Journal of Medicine 17:722-727.

Faure (1759). Recueil des pieces qui ont concouru pour le prix de L'Académie Royale de Chirurgie. Vol 8. Paris, P.Al Le Prieur.

Haygarth J (1800). Of the imagination, as a cause and as a cure of disorders of the body: exemplified by fictitious tractors, and epidemical convulsions. Bath: R. Crutwell.

Kaptchuk TJ (1998). Intentional ignorance: a history of blind assessment and placebo controls in medicine. *Bulletin of the History of Medicine* 72: 389-433.

Löhner G (1835), on behalf of a Society of truth-loving men. Die Homoöopathischen Kochsalzversuche zu Nürnberg [The homeopathic salt trials in Nuremberg].

Martini P (1932). *Methodenlehre der Therapeutischen Untersuchung*. Berlin:Springer.

Medical Research Council (1944). Clinical trial of patulin in the common cold. *Lancet* 2: 373-375 .

Medical Research Council (1948). Streptomycin treatment of pulmonary tuberculosis: a Medical Research Council investigation. *BMJ* 2: 769-782.

Medical Research Council (1950). Clinical trials of antihistaminic drugs in the prevention and treatment of the common cold. *BMJ* 2: 425-431.

Schulz KF, Chalmers I, Hayes RJ, Altman DG (1995). Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 273: 408-412.

Schulz KF, Chalmers I, Altman D (2002). The landscape and lexicon of blinding. *Annals of Internal Medicine* 136: 254-259.

Storke EF, Martin R, Rosenkrans EM, Ford J, Schloemilch A, McDermott GC, Carlson OW (1880). Final report of the Milwaukee test of the thirtieth dilution. *Homeopathic Times: A Monthly Journal of Medicine, Surgery and the Collateral Sciences* 7: 280-281.

[Home](#)

[Conteúdo](#)

[Envie comentários](#)

[Home](#)[Conteúdo](#)[jameslindlibrary.org](http://www.jameslindlibrary.org)

Interpretação de comparações sem vícios

Uma comparação controlada de tratamentos é aquela que [evita as comparações com viés](#). Isto envolve adotar medidas para minimizar os vieses decorrentes das [diferenças entre os pacientes comparados](#), e os [vieses decorrentes das diferenças na forma como os desfechos são avaliados](#).

Mesmo se esses vieses tiverem sido evitados, todavia, as interpretações das comparações sem viés nem sempre são claras. Por exemplo, foram consideradas as [diferenças entre os tratamentos planejados e os tratamentos recebidos](#) e o [efeito do acaso](#) foi considerado?

Algumas vezes, um novo estudo fornece uma evidência muito forte dos efeitos de um tratamento. Por exemplo, dezenas de milhares de pessoas participaram de um estudo extraordinário que revelou que um comprimido de aspirina podia reduzir substancialmente o risco de morte entre as pessoas que sofrem ataque cardíaco (ISIS-2 1988). É só muito raramente, entretanto, que um único estudo fornece uma evidência tão forte, assim é importante ao ler os relatórios da maioria dos estudos inquirir se a nova evidência foi integrada a [revisões sistemáticas de todas as outras evidências relevantes](#). Neste caso, foram tomadas medidas durante o processo de síntese para minimizar o impacto do [relatório com viés da evidência disponível](#) e da [seleção com viés da evidência disponível](#)? O potencial de [redução do efeito do acaso usando a meta-análise](#) foi considerado?

Cite as: Editorial commentary (2007). Interpretação de comparações sem vícios. The James Lind Library (www.jameslindlibrary.org).

Next essay: [Diferenças entre tratamento planejado e tratamento de fato realizado](#)

Select essay:

Referências

ISIS-2 Second International Study of Infarct Survival Collaborative Group (1988). Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17187 cases of suspected acute myocardial infarction: ISIS-2. *Lancet* 2: 349 60.

[Home](#)[Conteúdo](#)

[Envie comentários](#)

[Home](#)[Conteúdo](#)[jameslindlibrary.org](http://www.jameslindlibrary.org)

Interpretação de comparações sem viés:

Diferenças entre tratamento planejado e tratamento de fato realizado

Os experimentos controlados de tratamentos médicos devem ser planejados cuidadosamente. Os documentos que demonstram estes planos são referidos como protocolos, e, entre outras coisas, eles especificam os detalhes sobre os tratamentos que serão comparados. Porém, mesmo os melhores planos nem sempre funcionam conforme pretendido. Os tratamentos que de fato são recebidos pelos pacientes nos experimentos às vezes diferem daqueles que, conforme planejado, eles deviam ter recebido. Estes desvios da intenção precisam ser levados em conta na interpretação dos resultados das comparações de tratamentos.

Um dos motivos pelos quais os placebos foram introduzidos na evolução dos experimentos controlados de tratamentos médicos foi para reduzir o afastamento dos tratamentos pretendidos (Kaptchuk 1998). Mas as coisas podem sair do controle mesmo em estudos controlados por placebos. Durante a segunda guerra mundial, algumas pessoas que sofriam de resfriados receberam uma solução de antibiótico chamado patulina e foram comparadas com outras pessoas que receberam somente o líquido no qual o medicamento foi dissolvido ([MRC 1944](#)). A análise dos resultados não revelou quaisquer efeitos benéficos do medicamento, mas então surgiu uma preocupação de que o líquido usado para dissolver o medicamento poderia ter tirado o efeito do medicamento. Em outras palavras, mais de 1.000 pacientes podem ter participado de uma comparação de dois tratamentos inativos! Felizmente, experimentos confirmaram que a patulina usada no estudo foi eficaz, ainda que não tivesse nenhum efeito detectável nos resfriados (Chalmers and Clarke 2004)!

Os tratamentos recebidos podem diferenciar dos tratamentos pretendidos por várias razões. Por exemplo, os médicos podem decidir que o tratamento no qual alguns dos seus pacientes foram alocados em uma comparação de tratamentos formal não deve ser oferecido a eles; os pacientes podem rejeitar os tratamentos alocados para eles, ou não adotá-los como planejado; podem ser administradas doses do tratamento diferentes daquelas pretendidas; ou o fornecimento de um dos tratamentos pode se esgotar.

Por exemplo, quando surgiram diferenças nos resultados de tratamentos de leucemia aparentemente idênticos em crianças norte-americanas e britânicas, as investigações revelaram que o pior resultado na Grã-Bretanha refletiu a relutância entre os médicos britânicos de persistir com a quimioterapia quando os efeitos tóxicos desagradáveis do tratamento começaram a se desenvolver (Medical Research Council Working Party on Leukaemia in Children 1986).

Por esses motivos, as interpretações dos experimentos controlados devem considerar a possibilidade de que os tratamentos recebidos não eram os pretendidos. Se houve discrepâncias entre a intenção e a prática, é importante considerar as implicações para a interpretação da evidência.

Cite as: Editorial commentary (2007). Diferenças entre tratamento planejado e tratamento de fato realizado. The James Lind Library (www.jameslindlibrary.org).

Next essay: [Considerações sobre o efeito do acaso](#)

Select essay:

Referências

Chalmers I, Clarke M (2004). The 1944 Patulin Trial: the first properly controlled multicentre trial conducted under the aegis of the British Medical Research Council. *International Journal of Epidemiology* 32: 253-260.

Kaptchuk TJ (1998). Intentional ignorance: a history of blind assessment and placebo controls in medicine. *Bulletin of the History of Medicine* 72: 389-433.

Medical Research Council (1944). Clinical trial of patulin in the common cold. *Lancet* 2: 373-375 .

Medical Research Council Working Party on Leukaemia in Children (1986). Improvement in treatment for children with acute lymphoblastic leukaemia. *Lancet* 1: 408-11.

[Envie comentários](#)

[Home](#)[Conteúdo](#)jameslindlibrary.org

Interpretação de comparações sem viés:

Considerações sobre o efeito do acaso

Ao comparar dois tratamentos, quaisquer diferenças no desfecho podem ser simplesmente causadas pela obra do acaso. Por exemplo, pegue uma comparação de um novo tratamento com um tratamento padrão no qual 4 pessoas melhoraram com o primeiro e 6 pessoas melhoraram com o último. Seria obviamente errado concluir com segurança que o novo tratamento é pior do que o tratamento padrão: estes resultados podem simplesmente refletir o efeito do acaso. Se a comparação fosse repetida, os números dos pacientes que melhoraram poderiam se inverter (6 contra 4), ou serem iguais (5 contra 5), ou em alguma outra proporção.

Porém, se 40 pessoas melhoraram com o novo tratamento e 60 com o tratamento padrão, o acaso se torna uma explicação menos plausível para a diferença. E se 400 pessoas melhoraram com o novo tratamento e 600 com o tratamento padrão, seria evidente que o novo tratamento é de fato, muito provavelmente, pior do que o padrão. A maneira de reduzir a probabilidade de ser enganado pelo efeito do acaso nas comparações de tratamentos é assegurar-se de que os experimentos controlados incluam um número suficientemente grande de pessoas que experimentam o desfecho que se espera influenciar, melhorando ou piorando.

Em algumas circunstâncias um número muito grande de pessoas, milhares e às vezes dezenas de milhares, precisa participar de experimentos controlados para se obter avaliações confiáveis dos efeitos dos tratamentos. É necessário um grande número de pessoas, por exemplo, se os desfechos de interesse dos tratamentos são raros, por exemplo, ataques cardíacos e derrames entre mulheres de meia idade aparentemente saudáveis usando a terapia de reposição hormonal (TRH). Também é necessário um grande número de pessoas para que os efeitos moderados mas importantes dos tratamentos sejam detectados com segurança, por exemplo, uma redução de 20 por cento no risco de morte prematura entre pessoas que têm ataques cardíacos.

Para avaliar o papel do acaso nos resultados de experimentos controlados, os pesquisadores utilizam os "testes de significância estatística". Quando os estatísticos e outros recorrem às 'diferenças significantes' entre os tratamentos, eles estão geralmente se referindo à significância estatística. As diferenças estatisticamente significantes entre os tratamentos não têm necessariamente uma importância prática. Mas os testes de significância estatística são importantes, todavia, porque eles nos ajudam a evitar conclusões equivocadas de que existam diferenças reais nos tratamentos quando elas não existem, às vezes chamados de erros de Tipo I.

É importante levar em conta um número suficientemente grande de desfechos do tratamento para evitar um perigo muito mais comum, concluir que não existem diferenças entre os tratamentos quando de fato existem. Estes erros são às vezes chamados de erros de Tipo II. Thomas Graham Balfour sabia deste último perigo quando interpretou os resultados do seu teste das afirmações de que a beladona poderia prevenir os órfãos sob os seus cuidados de desenvolver escarlatina ([Balfour 1854](#)). Dois dos 76 garotos alocados para receber a beladona desenvolveram escarlatina comparado a 2 dos 75 garotos que não receberam o medicamento. Balfour observou que os "números são muito pequenos para justificar deduções quanto ao poder profilático da beladona". Se mais garotos tivessem desenvolvido escarlatina, Balfour poderia chegar a uma conclusão mais segura sobre os possíveis efeitos da beladona. Em vez disso, ele só observou que 4 casos de escarlatina entre 151 garotos era um número muito pequeno para se chegar a uma conclusão confiável.

Uma abordagem que reduz a probabilidade sermos enganados pelos efeitos do acaso implica em avaliar uma gama de diferenças do tratamento dentro da qual as diferenças reais podem residir ([Gavarret 1840](#); Huth 2006). Estas avaliações são conhecidas como intervalos de confiança. Conforme ilustrado no parágrafo de abertura deste artigo, repetir uma comparação de tratamentos pode gerar diversas avaliações dos efeitos diferenciais dos tratamentos nos desfechos, principalmente se as avaliações forem baseadas em um número pequeno de desfechos. Os intervalos de confiança levam em conta esta variação. Os intervalos de confiança são mais informativos do que simples testes de significância estatística, e deste modo mais úteis na redução da probabilidade de sermos enganados pelo efeito do acaso.

Os testes estatísticos e os intervalos de confiança, seja para análise de estudos individuais ou na [meta-análise](#) de um número de estudos isolados, porém semelhantes, nos ajudam a considerar o efeito do acaso e evitar a conclusão de que os efeitos e as diferenças dos tratamentos existem quando não existem, e de que não existem quando existem.

Cite as: Editorial commentary (2007). Considerações sobre o efeito do acaso. The James Lind Library (www.jameslindlibrary.org).

Show JLL records: illustrating [taking account of the play of chance](#)

Next essay: [Identificação dos efeitos imprevistos do tratamento](#)

Select essay:

Referências

Balfour TG (1854). Quoted in West C. Lectures on the Diseases of Infancy and Childhood. London, Longman, Brown, Green and Longmans, p 600.

Gavarret LDJ (1840). Principes généraux de statistique médicale: ou développement des règles qui doivent présider à son emploi. Paris: Bechet jeune & Labé.

Huth EJ (2006). Jules Gavarret's Principes Généraux de Statistique Médicale: a pioneering text on the statistical analysis of the results of treatments.

[Home](#)

[Conteúdo](#)

[Envie comentários](#)

Identificação dos efeitos imprevistos do tratamento

Os efeitos imprevistos de tratamentos só aparecerão quando novos tratamentos forem introduzidos mais amplamente. Experimentos iniciais, por exemplo, aqueles exigidos para autorizar novos medicamentos, abrangem não mais do que algumas centenas ou algumas milhares de pessoas tratadas por alguns meses. Apenas os efeitos imprevistos relativamente frequentes e de curto prazo podem ser obtidos neste estágio.

Os efeitos raros dos tratamentos, ou aqueles que levam algum tempo para se desenvolver, não serão descobertos até que os testes do tratamento tenham durado o bastante ou até que haja um uso mais frequente dos tratamentos. Além disso, os novos tratamentos serão, quase sempre, usados em pessoas que possam se diferenciar de forma significativa daquelas que participaram dos experimentos originais. Elas podem ser mais velhas ou mais jovens, de sexos diferentes, estar mais ou menos doentes, viver em circunstâncias diferentes, ou sofrer de outros problemas de saúde, além da condição para a qual o tratamento é orientado. Estas diferenças podem modificar os efeitos dos tratamentos, e podem surgir efeitos novos e imprevistos (ver a edição especial da [BMJ de 3 de julho de 2004](#)).



A detecção e a verificação de efeitos imprevistos, sejam [adversos](#) ou [benéficos](#), geralmente ocorrem de forma bem diferente dos métodos usados para avaliar os efeitos esperados dos novos tratamentos. É comum os efeitos imprevistos dos tratamentos serem suspeitados inicialmente pelos profissionais da saúde ou pelos pacientes. Identificar quais dentre as desconfianças iniciais refletem os efeitos reais dos tratamentos é um desafio, que se tornou familiar aos leitores dos artigos anteriores desta série, ou seja, para evitar ser enganado por [vieses](#) e pelo [efeito do acaso](#).

Se o efeito imprevisto de um tratamento for muito evidente e ocorrer com muita frequência após o uso do tratamento, ele poderá ser notado espontaneamente pelos profissionais da saúde ou pelos pacientes. Por exemplo, quase não se ouve falar de bebês nascidos sem os membros, assim quando na década de 60 houve um aumento repentino no número de nascimentos destes bebês, aumentou também o interesse sobre o assunto. Todas as mães de tais bebês tinham usado um medicamento anti-náusea, a [talidomida](#), recentemente comercializado, prescrito durante o início da gestação, deste modo esta era a causa provável, e poucas avaliações adicionais foram necessárias. Os efeitos benéficos imprevistos dos medicamentos são frequentemente detectados de maneiras semelhantes, por exemplo, quando se descobriu que um medicamento usado para tratar a psicose também baixava o colesterol (Goodwin 1991).

Quando relações tão impressionantes são observadas, elas geralmente acabam sendo confirmadas como efeitos imprevistos reais do tratamento ([Venning 1982](#)). Todavia, muitas desconfianças sobre os efeitos imprevistos dos tratamentos são baseadas em evidências muito menos convincentes. Assim como os experimentos criados para detectar os efeitos esperados dos tratamentos, planejar experimentos para confirmar ou descartar suspeitas de efeitos imprevistos não tão evidentes, implica em [evitar comparações com viés](#).

Os estudos para testar se as suspeitas de efeitos imprevistos do tratamento são reais devem observar o princípio de comparação de 'semelhante com semelhante'. A alocação randomizada para os tratamentos é a maneira ideal para executar esta tarefa. Muito raramente, entretanto, as suspeitas dos efeitos dos tratamentos podem ser investigadas por uma análise mais detalhada ou pelo acompanhamento das pessoas que foram alocadas aleatoriamente nos tratamentos antes de eles serem administrados (Hemminki and McPherson 1997). O desafio é portanto formar grupos imparciais de comparação de outras maneiras, quase sempre usando as informações coletadas de modo rotineiro durante a assistência médica.

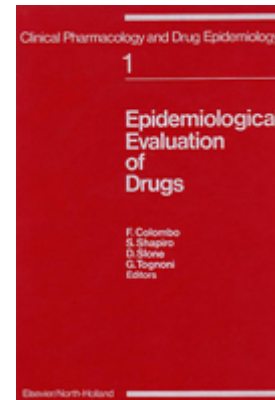
Nestes estudos, o fato de os efeitos não terem sido antecipados no momento em que as decisões do tratamento forem tomadas é, na verdade, benéfico. Pois isto significa que o risco da condição suspeita não poderia ter sido considerado no momento em que as pessoas estavam sendo selecionadas para o tratamento: o efeito imprevisto é geralmente uma condição ou doença diferente da condição ou doença para a qual o tratamento foi prescrito (Vandenbroucke 2004a).

Por exemplo, quando a terapia de reposição hormonal (TRH) foi introduzida para tratar os sintomas da menopausa, o risco de uma mulher desenvolver trombose venosa provavelmente não tenha sido levado em conta, porque a maioria dos médicos e das mulheres pensavam que isto era irrelevante. Não havia, portanto, nenhuma razão para esperar que as mulheres que receberam a TRH tivessem um risco de desenvolver uma trombose venosa diferente daquelas que não receberam o medicamento. Deste modo, estabeleceu-se a base para os experimentos controlados, e estes mostraram que a TRH aumenta o risco da trombose venosa.

Quando a suspeita de um efeito imprevisto está relacionada a um tratamento de um problema de saúde comum (como ataque cardíaco), mas não ocorre com muita frequência com o novo tratamento (ou não é completamente

aliviado por ele), é necessária uma vigilância em larga escala das pessoas que estão recebendo o tratamento para se detectar o efeito imprevisto. Por exemplo, embora alguma pessoas acreditassem que a [aspirina](#) podia reduzir o risco de ataque cardíaco, e no final dos anos 60, tivessem começado a fazer experimentos controlados dessa teoria ([Elwood et al. 1974](#)), a maioria das pessoas pensava que esta teoria era bastante improvável. O avanço veio quando foi realizado um grande estudo para detectar os efeitos adversos imprevistos de medicamentos: os pesquisadores observaram que as pessoas que davam entrada em hospitais com ataques cardíacos tinham menos probabilidade de terem tomado aspirina recentemente do que pacientes aparentemente semelhantes ([Boston Collaborative Drug Surveillance Group 1974](#)). Estas conclusões eram consistentes com aquelas de um experimento controlado, no qual as pessoas tinham sido alocadas para receber aspirina ou não após o ataque cardíaco. Os dois relatórios foram publicados consecutivamente na mesma edição da Revista Britânica de Medicina (British Medical Journal) (BMJ 1974).

Os princípios básicos para a detecção e a investigação dos efeitos imprevistos dos tratamentos foram estipulados pela primeira vez no final dos anos 70 ([Jick 1977](#); [Colombo et al. 1977](#)). Eles se inspiraram na experiência coletiva da investigação de efeitos imprevistos que se acumulou após o desastre da [talidomida](#). Os requisitos para um importante tipo de pesquisa, estudos de caso-controle dos possíveis efeitos adversos, foram colocados no papel baseados nas experiências de pesquisadores de Boston e Oxford ([Jick and Vessey 1978](#)). Com muitos a introdução de muitos tratamentos eficazes desde aquele tempo, este aspecto dos experimentos controlados de tratamentos permanece tão desafiador e importante hoje quanto era na época (Vandenbroucke 2004b; Vandenbroucke 2006; Papanikolaou et al. 2006).



Conforme enfatizado nos artigos anteriores desta série, é importante reconhecer que relatórios individuais que sugerem ou descartam as suspeitas sobre os efeitos imprevistos de tratamentos podem ser equivocados. Como com todos os outros experimentos controlados, os possíveis efeitos imprevistos do tratamento devem ser investigados através de [revisões sistemáticas](#) de todas as evidências relevantes, como aquelas que confirmaram a relação entre a TRH e o ataque cardíaco, o derrame e o câncer de mama (Hemminki and McPherson 1997; Collaborative Group on Hormonal Factors in Breast Cancer 1997).

Cite como: Editorial commentary (2007). Identificação dos efeitos imprevistos do tratamento. The James Lind Library (www.jameslindlibrary.org).

Mostre registros de JL: illustrating [identification of unanticipated effects](#)

Essay seguinte: [Revisões sistemáticas das evidências relevantes](#)

Selecione o essay:

Referências

Boston Collaborative Drug Surveillance Group (1974). Regular aspirin intake and acute myocardial infarction. *BMJ* 1: 440-443.

Collaborative Group on Hormonal Factors in Breast Cancer (1997). Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. *Lancet* 350: 1047-1059 .

Colombo F, Shapiro S, Slone D, Tognoni G, eds (1977). *Epidemiological Evaluation of Drugs*. Amsterdam: Elsevier/ North Holland Biomedical Press, 1977.

Elwood PC, Cochrane AL, Burr ML, Sweetnam PM, Williams G, Welsby E, Hughes SJ, Renton R (1974). A randomised controlled trial of acetyl salicylic acid in the secondary prevention of mortality from myocardial infarction. *BMJ* 1: 436-440.

Goodwin JS (1991). The empirical basis for the discovery of new therapies. *Perspectives in Biology and Medicine* 35: 20-36.

Hemminki E, McPherson K (1997). Impact of postmenopausal hormone therapy on cardiovascular events and cancer: pooled data from clinical trials. *BMJ*; 315: 149-153.

Jick H (1977). The discovery of drug-induced illness. *New England Journal of Medicine* 296: 481-485.

Jick H, Vessey M (1978). Case-control studies in the evaluation of drug-induced illness. *American Journal of Epidemiology* 107: 1-7.

Papanikolaou PN, Christidi GD, Ioannidis JPA (2006). Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies. *CMAJ* 174: 635-641.

Vandenbroucke JP (2004a). When are observational studies as credible as randomised trials? *Lancet* 363: 1728-1731 .

Vandenbroucke JP (2004b). Benefits and harms of drug treatments. *BMJ* 329: 2-3.

Vandenbroucke JP (2006). What is the best evidence for determining harms of medical treatment? *CMAJ* 174: 645-646.

Venning GR (1982). Validity of anecdotal reports of suspected adverse drug reactions: the problem of false alarms. *BMJ* 284: 249-254.

[Página inicial](#)

[Conteúdo](#)



Translated from the original English by the Pan American Health Organization.
All rights reserved. www.paho.org

[Envie comentários](#)

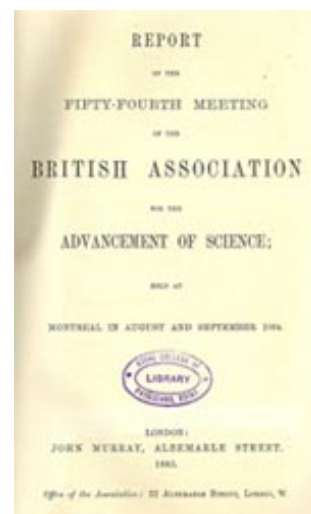
[Home](#)[Conteúdo](#)jameslindlibrary.org

Revisões sistemáticas das evidências relevantes

Um dos pioneiros do século vinte em experimentos controlados de tratamentos, [Austin Bradford Hill](#), observou que os leitores de relatórios de pesquisa querem respostas para quatro perguntas: 'Por que você começou?', 'O que você fez?', 'O que você encontrou?' e 'O que isso significa afinal?' (Hill 1965). O qualidade da resposta para a última pergunta de Hill é muito importante porque este é o elemento de um relatório de pesquisa que mais pode influenciar as escolhas e as decisões sobre os tratamentos.

Só muito raramente um único experimento controlado de um tratamento produz evidências suficientemente fortes para fornecer uma resposta segura à pergunta 'O que isso significa?'. Um experimento controlado de um tratamento é geralmente um de vários experimentos que abordam a mesma questão. Para uma resposta confiável à pergunta 'O que isso significa?', então, é importante interpretar a evidência de um experimento controlado específico no contexto de uma avaliação cuidadosa de todas as evidências dos experimentos controlados que abordaram a questão envolvida.

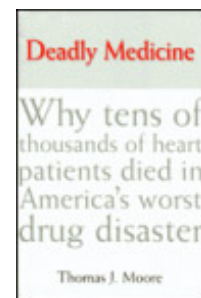
O presidente da Associação Britânica para o Avanço da Ciência expressou a necessidade de observar este princípio há mais de um século:



"Se, como algumas vezes se supõe, a ciência consistisse em nada além da acumulação laboriosa de fatos, logo viria uma pausa, esmagada, por assim dizer, pelo seu próprio peso.... Dois processos estão deste modo em ação lado a lado, a recepção de novos materiais e a digestão e a assimilação dos antigos...O trabalho que merece, mas que receio nem sempre receba, o maior crédito é aquele em que a descoberta e a explicação andam de mãos dadas, no qual sejam apresentados não somente os novos fatos, mas sua relação com os antigos". ([Rayleigh 1885](#))

Talvez seja porque a aplicação deste princípio na prática ainda atraia pouco crédito dentro do ambiente acadêmico, que poucos relatórios de experimentos controlados de tratamentos discutam os seus resultados no contexto de uma avaliação sistemática de todas as outras evidências relevantes (Clarke et al. 2002). Como resultado, é geralmente difícil para os leitores obterem uma resposta confiável à pergunta 'O que isso significa?' dos relatórios de novas pesquisas.

Conforme observado em um artigo explicativo anterior, iniciar novos experimentos de tratamentos médicos sem primeiro revisar sistematicamente o que pode ser aprendido a partir de pesquisas já existentes é perigoso, imprudente e antiético (ver [Por que as comparações devem tratar de incertezas legítimas](#)). Relatar os resultados de novos experimentos sem interpretar novas evidências à luz de avaliações sistemáticas de outras evidências relevantes também é perigoso porque resulta em atrasos na identificação tanto de tratamentos benéficos quanto maléficis (Antman et al. 1992). Por exemplo, entre a década de 60 e o início da década de 90, foram feitos mais de 50 experimentos controlados de medicamentos para reduzir as anormalidades do batimento cardíaco em pessoas que tiveram ataques cardíacos antes que se percebesse que estes medicamentos estavam matando as pessoas. Se cada relatório tivesse avaliado os resultados dos novos experimentos no contexto de todas as evidências relevantes, os efeitos letais do medicamento poderiam ter sido identificados uma década mais cedo, e muitas mortes prematuras desnecessárias poderiam ter sido evitadas.



Em uma era de publicação eletrônica deveria ser possível lidar com as limitações encontradas na maioria dos relatórios de novas pesquisas (Chalmers and Altman 1999; Smith and Chalmers 2001). Entretanto, ao invés de

fundamentar as conclusões sobre os tratamentos em um ou alguns poucos estudos individuais, os usuários de evidências de pesquisas estão cada vez mais se baseando em informações confiáveis para as revisões [sistemáticas e atualizadas de todas as evidências relevantes e confiáveis](#), porque estas são cada vez mais reconhecidas como sendo as que fornecem a melhor base para conclusões sobre os efeitos dos tratamentos médicos.

Da mesma forma que é importante tomar medidas para evitar ser enganado por [vieses](#) e pelo [efeito do acaso](#) no planejamento, na análise e na interpretação de experimentos controlados de tratamentos, assim também devemos tomar medidas no planejamento, execução, análise e interpretação das revisões sistemáticas. Isto envolve:

- especificar a questão a ser tratada pela revisão sistemática
- definir o critério de elegibilidade dos estudos a serem incluídos
- identificar (todos) os possíveis estudos elegíveis
- aplicar o critério de elegibilidade de maneira que limite o viés
- formar uma proporção mais alta possível das informações relevantes dos estudos
- analisar estas informações, se apropriado e possível, usando a meta-análise e uma variedade de análises
- preparar um relatório estruturado

Uma manifestação do aumento do reconhecimento da importância crucial das revisões sistemáticas para a avaliação dos efeitos dos tratamentos é a rápida evolução dos métodos para melhorar a confiabilidade das revisões em si. A primeira edição de um livro intitulado Systematic Reviews (Revisões Sistemáticas) tinha menos de 100 páginas ([Chalmers and Altman 1995](#)): apenas seis anos mais tarde, a segunda edição trazia aproximadamente 500 páginas e incluía estratégias muito avançadas para aumentar as informações obtidas nas pesquisas (Egger et al. 2001).

Há atualmente desenvolvimentos importantes nos métodos de preparo das revisões sistemáticas, incluindo aqueles necessários para identificar os efeitos imprevistos dos tratamentos (Glasziou et al. 2004) e para incorporar os resultados da pesquisa que descrevem e analisam as experiências das pessoas que ministraram e que receberam os tratamentos (Thomas 2004). À medida que forem surgindo, os materiais relevantes serão acrescentados à James Lind Library.

Cite as: Editorial commentary (2007). Revisões sistemáticas das evidências relevantes. The James Lind Library (www.jameslindlibrary.org).

Show JLL records: illustrating [systematic review](#)

Next essay: [Como lidar com o vício da evidência disponível](#)

Select essay:

Referências

Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC (1992). A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. JAMA 268:240-48.

Chalmers I, Altman DG (1995). Systematic Reviews. London: BMJ Publications.

Chalmers I, Altman DG (1999). How can medical journals help prevent poor medical research? Some opportunities presented by electronic publishing. Lancet 353:490-493 .

Egger M, Davey Smith G, Altman D (2001). Systematic Reviews in Health Care: meta-analysis in context. 2nd Edition of Systematic Reviews. London: BMJ Books.

Glasziou P, Vandenbroucke J, Chalmers I (2004). Assessing the quality of research BMJ 328:39-41.

Hill AB (1965). Cited in 'The reasons for writing'. BMJ 4:870.

Rayleigh (1885). Address by the Rt. Hon. Lord Rayleigh. In: Report of the fifty-fourth meeting of the British Association for the Advancement of Science; held at Montreal in August and September 1884, London: John Murray.

Smith R, Chalmers I (2001). Britain's gift: a 'Medline' of synthesized evidence. BMJ 323:1437-1438.

Thomas J, Harden A, Oakley A, Oliver S, Sutcliffe K, Rees R, Brunton G, Kavanagh J (2004). Integrating qualitative

[Envie comentários](#)

Revisões sistemáticas das evidências relevantes:

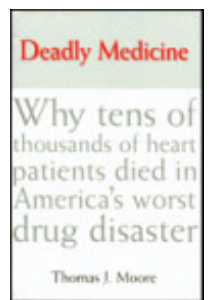
Como lidar com o vício da evidência disponível

[Evitar as comparações com viés](#) envolve identificar e considerar todas as evidências relevantes e confiáveis nas revisões sistemáticas. Isto é um desafio em muitos aspectos, particularmente porque algumas evidências pertinentes não são publicadas devido a decisões com viés sobre quais resultados da pesquisa são apresentados e aceitos para publicação. Os estudos que produzem resultados "decepcionantes" ou "negativos" têm menos probabilidade de serem relatados do que outros. Isto é, geralmente, chamado de 'viés de publicação' ou 'viés de informação'.

Estes vieses de informação têm sido reconhecidos há séculos (Dickersin 2004a). Em 1792, por exemplo, James Ferriar enfatizou a importância de se registrar os fracassos tanto quanto os sucessos dos tratamentos (Ferriar 1792). Este princípio foi reiterado em um editorial publicado na Revista de Medicina e Cirurgia de Boston (Boston Medical and Surgical Journal) cerca de um século mais tarde ([Editorial 1909](#)).

Existe agora um grande conjunto de evidências confirmando que o viés de informação é um problema importante. Há também evidência de que o viés de informação resulta principalmente do fato de os pesquisadores não relatarem por escrito ou submeterem relatórios de pesquisa para publicação, não por causa da rejeição com viés dos relatórios submetidos por editores de revistas (Dickersin 2004b). Pesquisas recentes também revelaram um outro problema: se as estimativas dos efeitos dos tratamentos de alguns desfechos estudados não sustentam as conclusões dos pesquisadores, estes dados, às vezes, também não são relatados (Chan et al. 2004).

Por exemplo, se todos os estudos sobre os efeitos da administração de medicamentos para reduzir as anormalidades do batimento cardíaco em pacientes que tiveram ataques cardíacos tivessem sido relatados, dezenas de milhares de mortes causadas por estes medicamentos poderiam ter sido evitadas. Em 1993, o Dr. Cowley e seus colegas demonstraram como um estudo não publicado feito 13 anos antes poderia ter "fornecido um alerta precoce de problemas que estavam por vir". Nove pacientes tinham morrido entre os 49 que receberam o medicamento anti arritmia (lorcainida) comparado com somente um paciente entre um número semelhante de pessoas que receberam placebos. "Quando realizamos nosso estudo em 1980", eles relataram, "pensamos que a elevada taxa de mortalidade era um efeito do acaso...O desenvolvimento da lorcainida foi abandonado por motivos comerciais, e esse estudo, portanto, nunca foi publicado; agora é um bom exemplo de 'viés de publicação'" (Cowley et al. 1993).



Vieses de informação tendem a levar a conclusões de que os tratamentos médicos são mais eficazes do que eles são na verdade. Eles podem, portanto, resultar em sofrimento e morte desnecessários, e em recursos desperdiçados em tratamentos ineficazes e perigosos (Chalmers 2004). As pessoas que concordam com o pedido dos pesquisadores para que participem nos experimentos de tratamentos supõem que a sua participação causará um aumento do conhecimento. Este contrato implícito entre os pesquisadores e participantes da pesquisa é quebrado pelos pesquisadores que não tornam públicos os resultados da pesquisa.

A subnotificação com viés da pesquisa é considerada má conduta científica e é antiética (Chalmers 1990). O relato seletivo dos estudos patrocinados pela indústria farmacêutica é um outro problema (Melander et al. 2003), ainda que o problema não esteja limitado àqueles com interesses comerciais escusos. Os comitês de ética na pesquisa, os especialistas em ética médica e as instituições de pesquisa até agora não fizeram o bastante para proteger os pacientes e o público dos efeitos adversos dos vieses de informação (Savulescu et al. 1996). Os experimentos controlados de tratamentos, especialmente aqueles tratamentos nos quais há um interesse comercial, permanecerão comprometidos enquanto estas formas de má-conduta da pesquisa forem toleradas pelos governos e por outros que deveriam proteger os interesses públicos.

A Organização Mundial de Saúde propôs soluções para tratar do problema do viés da pesquisa e da publicação (ou disseminação) não identificáveis: primeiramente, deve-se estabelecer padrões para o registro e o intercâmbio de dados para o registro de estudos. Em segundo lugar, propõe-se o registro dos protocolos de pesquisa em bases de dados que observem os padrões acima, antes de iniciar o recrutamento de pacientes. Finalmente, propõe-se a implementação de um portal de acesso livre (www.who.int/ictpr), que confere o dados de todos os registros, permitindo que as pessoas se informem sobre protocolos de pesquisas que estão começando, que estão em andamento e que já acabaram.

Todos devemos seguir o exemplo dado pela Organização Mundial de Saúde para reduzir os vieses de informação,

exigindo o registro de todos os experimentos controlados de tratamentos no seu início e insistindo que os seus resultados sejam publicados.

Cite as: Editorial commentary (2007). Como lidar com o vício da evidência disponível. The James Lind Library (www.jameslindlibrary.org).

Show JLL records: illustrating [reporting bias](#)

Other material: view [Monty Python's take](#) on reporting bias - YouTube video

Next essay: [Precaução com o vício da evidência disponível](#)

Select essay:

Referências

Chalmers I (1990). Under-reporting research is scientific misconduct. *JAMA* 263:1405-1408.

Chalmers I (2004). In the dark: drug companies should be forced to publish all the results of clinical trials. *New Scientist* 181:19.

Chan A-W, Hróbjartsson A, Haahr M, Gøtzsche PC, Altman DG (2004). Empirical evidence for selective reporting of outcomes in randomized trials: Comparison of protocols to publications. *JAMA* 291:2457-2465.

Cowley AJ, Skene A, Stainer, Hampton JR (1993). The effect of lorcinide on arrhythmias and survival in patients with acute myocardial infarction. *International Journal of Cardiology* 40:161-166.

Dickersin K (2004a). Publication bias: recognising the problem, understanding its origins and scope, and preventing harm. In: Rothstein H, Sutton A, Borenstein M, eds. *Handbook of publication bias*. New York: Wiley.

Dickersin K (2004b). How important is publication bias? A synthesis of available data. *AIDS Educ Prev* 1997;9 (1 Suppl): 15-21.

Editorial (1909). The reporting of unsuccessful cases. *Boston Medical and Surgical Journal* 161:263-264.

Ferriar J (1792). *Medical histories and reflexions*. Vol 1. London: Cadell and Davies, 1792.

Melander H, Ahlqvist-Rastad J, Meijer G, Beermann B (2003). Evidence b(i)ased medicine - selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. *BMJ* 326:1171-3.

Savulescu J, Chalmers I, Blunt J (1996). Are research ethics committees behaving unethically? Some suggestions for improving performance and accountability. *BMJ* 313:1390-1393.

[Home](#)

[Conteúdo](#)

[Envie comentários](#)

[Home](#)[Conteúdo](#)[jameslindlibrary.org](http://www.jameslindlibrary.org)

Revisões sistemáticas das evidências relevantes:

Precaução com o vício da evidência disponível

Os vieses podem distorcer os experimentos de tratamentos médicos e levar a conclusões errôneas. Eles também podem distorcer as revisões das evidências. O planejamento das revisões sistemáticas devem ser exibidos em protocolos, como aqueles publicados pela [Colaboração Cochrane](#), deixando claro quais medidas serão tomadas para reduzir os vieses.

Isto inclui especificar claramente:

- qual questão sobre os tratamentos será tratada na revisão;
- os critérios que qualificarão um estudo para inclusão;
- as estratégias que serão usadas para procurar estudos potencialmente elegíveis; e
- as medidas que serão tomadas para minimizar os vieses na seleção dos estudos e dos dados para inclusão na revisão (Berlin 1997).

Diferentes revisões sistemáticas tratando o que parece ser a mesma questão sobre os efeitos dos tratamentos médicos freqüentemente chegam a conclusões diferentes. Às vezes isto acontece porque as questões tratadas possuem pequenas diferenças. Às vezes apresentam diferenças nos materiais e métodos usados pelos revisores, e nestas circunstâncias é importante julgar quais das revisões são mais propensas a terem reduzido os vieses de alocação com maior êxito.

Também vale a pena considerar se os revisores têm outros interesses que possam afetar a realização ou a interpretação da sua revisão. Por exemplo, pessoas associadas aos fabricantes do óleo de primula da noite revisaram os efeitos do medicamento em eczemas (Morse et al. 1989). Eles chegaram a uma conclusão muito mais interessante sobre o valor do medicamento do que uma revisão feita por investigadores sem nenhum interesse comercial, que incluíram os resultados de estudos não publicados em sua avaliação (Williams 2003).

Não são somente os interesses comerciais que podem levar à seleção com viés de evidências disponíveis para inclusão em revisões. Todos nós temos preconceitos que podem levar à seleção com viés da evidência, e pesquisadores, profissionais da saúde, pacientes e outras pessoas que possam estar avaliando os efeitos dos tratamentos não estão imunes. Conflitos de interesse foram reconhecidos como importantes, e algumas medidas estão sendo tomadas para levá-los em conta.

Cite as: Editorial commentary (2007). Precaução contra a seleção com viés a partir da evidência disponível. The James Lind Library (www.jameslindlibrary.org).

Show JLL records: illustrating [reporting bias](#)

Other material: view [Monty Python's take](#) on reporting bias - YouTube video

Next essay: [Redução do efeito do acaso usando a meta-análise](#)

Select essay:

Referências

Berlin JA (1997). Does blinding of readers affect the results of meta-analyses? University of Pennsylvania Meta-analysis Blinding Study Group. *Lancet* 350:185-186 .

Morse PF, Horrobin DF, Manku MS, Stewart JC, Allen R, Littlewood S, Wright S, Burton J, Gould DJ, Holt PJ, et al (1989). Meta-analysis of placebo-controlled studies of the efficacy of Epogam in the treatment of atopic eczema. Relationship between plasma essential fatty acid changes and clinical response. *British Journal of Dermatology* 121:75-90.

Sackett DL, Oxman AD (2003). HARLOT plc: an amalgamation of the world's two oldest professions *BMJ* 327:1442-1445.

Williams HC (2003). Evening primrose oil for atopic dermatitis. BMJ 327: 1358-1359.

[Home](#)

[Conteúdo](#)

[Envie comentários](#)

[Home](#)[Conteúdo](#)jameslindlibrary.org

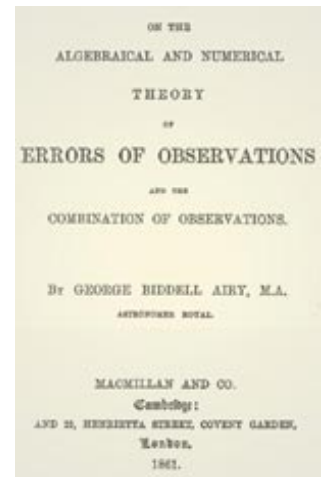
Revisões sistemáticas das evidências relevantes:

Redução do efeito do acaso usando a meta-análise

[Revisões sistemáticas](#) de todas as evidências relevantes e confiáveis são necessárias para a realização de experimentos controlados de tratamentos médicos. Para evitar conclusões equivocadas sobre os efeitos dos tratamentos, as pessoas que preparam as revisões sistemáticas devem tomar algumas medidas para evitar [vieses](#) de vários tipos, por exemplo, [considerando todas as evidências relevantes](#) e [evitando a seleção com vies das evidências disponíveis](#).

Embora se deva tomar cuidado para minimizar os vieses nas revisões, conclusões equivocadas sobre os efeitos dos tratamentos também podem ser [causadas pelo acaso](#). Discutir estudos distintos mas semelhantes um de cada vez, nas revisões sistemáticas, pode deixar uma impressão confusa por causa do efeito do acaso. Se for possível e apropriado, este problema pode ser reduzido pela combinação dos dados de todos os estudos relevantes, usando um procedimento estatístico agora conhecido como "meta-análise".

A maioria das técnicas estatísticas usadas hoje na meta-análise tem origem nos trabalhos do matemático alemão Karl Gauss e do matemático francês Pierre-Simon Laplace durante a primeira metade do século XIX. Um dos campos em que seus métodos encontraram aplicação prática foi na astronomia: medir a posição das estrelas em várias ocasiões, não raro, resultava em estimativas um pouco diferentes, assim, eram necessárias técnicas para combinar as estimativas para produzir uma média derivada da combinação dos resultados. Em 1861, o astrônomo real britânico George Airy, publicou um 'manual' para astrônomos ([Airy 1861](#)) no qual ele descrevia os métodos usados para este processo de síntese quantitativa. Cerca de um século depois, um cientista social norte-americano, Gene Glass, nomeou o processo de 'meta-análise' (Glass 1976).



Um dos primeiros exemplos médicos da meta-análise foi publicado na Revista Britânica de Medicina (British Medical Journal) em 1904 por Karl Pearson ([Pearson 1904](#); O'Rourke 2006), a quem o governo pediu que revisasse a evidência dos efeitos de uma vacina contra tifoide. Embora os métodos da meta-análise tenham sido desenvolvidos por estatísticos durante os 70 anos subsequentes, foi só na década de 70 que estes métodos começaram a ser aplicados mais amplamente, inicialmente por cientistas sociais (Glass 1976) e depois por pesquisadores médicos (Stjernswärd J 1974; Stjernswärd et al. 1976; Cochran et al. 1977; Chalmers et al. 1977; Chalmers 1979; Editorial 1980).

O método de meta-análise pode ser demonstrado usando o logotipo da [Colaboração Cochrane](#). O logotipo ilustra uma meta-análise dos dados de sete experimentos controlados. Cada linha horizontal representa os resultados de um teste (quanto mais curta a linha, mais certo é o resultado) e o diamante representa os seus resultados combinados. A linha vertical indica a posição ao redor da qual as linhas horizontais se agrupariam se os dois tratamentos comparados nos estudos tivessem efeitos semelhantes; se uma linha horizontal atravessasse a linha vertical, significaria que aquele experimento em particular não encontrou nenhuma diferença clara ("estatisticamente significativa") entre os tratamentos. Quando as linhas horizontais individuais atravessam a linha vertical 'sem diferença' significa que o tratamento pode tanto aumentar quanto diminuir as mortes de bebês. Juntas, todavia, as linhas horizontais tendem a cair no lado benéfico (esquerdo) da linha 'sem diferença'. O diamante representa os resultados combinados destes testes, gerados usando o processo estatístico de meta-análise. A posição do diamante nitidamente à esquerda da linha 'sem diferença' indica que o tratamento é benéfico.



Este diagrama mostra os resultados de uma revisão sistemática de experimentos controlados de um processo barato e curto de um medicamento esteróide administrado a mulheres que corriam o risco de dar à luz prematuramente. O primeiro destes testes foi relatado em 1972. O diagrama resume a evidência que teria sido revelada se os experimentos disponíveis tivessem sido revisados sistematicamente uma década mais tarde, em 1981: há uma forte indicação de que os esteróides reduzem o risco de bebês morrerem de complicações causadas

pela imaturidade. Até 1991, mais sete ensaios tinham sido relatados, e a imagem no logotipo tinha ficado ainda mais forte.

Nenhuma revisão sistemática destes estudos foi publicada até 1989 (Crowley 1989), assim a maioria dos obstetras e das mulheres grávidas não sabia que o tratamento era tão eficaz. Afinal, alguns experimentos não mostraram um benefício "estatisticamente significativo", e talvez apenas estes experimentos tinham sido observados. Como nenhuma revisão sistemática tinha sido feita, dezenas de milhares de bebês prematuros sofreram e morreram desnecessariamente, além do desperdício de recursos em pesquisas desnecessárias. Este é apenas um dos muitos exemplos de custos de vidas humanas que podem resultar do fracasso em avaliar os efeitos de tratamentos em [revisões sistemáticas](#) atualizadas dos experimentos controlados, usando a meta-análise para reduzir a probabilidade de que o [efeito do acaso](#) seja enganoso.

No fim do século XX, aceitou-se amplamente que a meta-análise era um elemento importante dos experimentos controlados, e que ela ajudava a evitar conclusões incorretas de que os tratamentos não tinham efeitos quando eles eram, na verdade, benéficos ou prejudiciais.

Cite as: Editorial commentary (2007). Redução do efeito do acaso usando a meta-análise. The James Lind Library (www.jameslindlibrary.org).

Show JLL records: illustrating [meta-analysis](#)

Next essay: [Avaliações sistemáticas atualizadas das evidências relevantes e confiáveis](#)

Select essay:

Referências

Airy GB (1861). On the algebraical and numerical theory of errors of observations and the combination of observations. London: Macmillan.

Chalmers I (1979). Randomized controlled trials of fetal monitoring 1973-1977. In: Thalhammer O, Baumgarten K, Pollak A, eds. Perinatal Medicine. Stuttgart: Georg Thieme, 260-265.

Chalmers TC, Matta RJ, Smith H, Kunzler A-M. (1977). Evidence favoring the use of anticoagulants in the hospital phase of acute myocardial infarction. New England Journal of Medicine 297:1091-1096.

Crowley P (1989). Promoting pulmonary maturity. In: Chalmers I, Enkin M, Keirse MJNC, eds. Effective care in pregnancy and childbirth. Oxford: Oxford University Press, pp 746-762.

Editorial (1980). Aspirin after myocardial infarction. Lancet 1:1172-3 .

Glass GV (1976). Primary, secondary and meta-analysis of research. Educational Researcher 10, 3-8.

O'Rourke K (2006). An historical perspective on meta-analysis: dealing quantitatively with varying study results. The James Lind Library.

Pearson K (1904). Report on certain enteric fever inoculation statistics. BMJ 3:1243-1246.

Stjernswärd J (1974). Decreased survival related to irradiation postoperatively in early operable breast cancer. Lancet 2:1285-1286.

Stjernswärd J, Muenz LR, von Essen CF (1976). Postoperative radiotherapy and breast cancer. Lancet 1:749.

[Home](#)

[Conteúdo](#)

[Envie comentários](#)

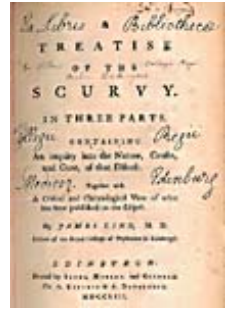
[Home](#)[Conteúdo](#)jameslindlibrary.org

Avaliações sistemáticas atualizadas das evidências relevantes e confiáveis

Experimentos controlados de tratamentos no atendimento à saúde

Os resultados dos experimentos controlados de tratamentos médicos são, muito raramente, ajustados sistematicamente no contexto de outros estudos semelhantes, usando métodos para reduzir os [vieses](#) e o [efeito do acaso](#). A falha em fazer revisões sistemáticas da pesquisa sobre os efeitos de tratamentos resultou em muito sofrimento que poderia ser evitado. Os [experimentos controlados de tratamentos na atenção à saúde](#) envolvem a preparação de revisões sistemáticas imparciais de todos os estudos de pesquisa relevantes e confiáveis dos tratamentos avaliados.

Existem alguns exemplos deste processo há mais de 200 anos atrás. Em 1753, por exemplo, em sua revisão do grande número de relatórios sobre a prevenção e o tratamento do escorbuto, James Lind observou:



"Como não é fácil destruir os preconceitos ... tornou-se necessário exibir uma visão imparcial e plena do que foi publicado até aqui sobre o escorbuto ... Na verdade, antes do assunto ser definido sob uma perspectiva nítida e adequada, foi necessário remover muito lixo". ([Lind 1753](#))

Revisões sistemáticas de todas as pesquisas relevantes que abordam as questões sobre os efeitos de tratamentos são, cada vez mais, vistas como fornecedoras de uma base mais segura para as conclusões sobre os efeitos de tratamentos. Às vezes, as revisões sistemáticas mostrarão que não existe nenhuma evidência confiável, e esta é uma das suas funções mais importantes. Da mesma forma, as revisões sistemáticas podem, às vezes, confirmar que a evidência confiável está limitada a um único estudo, e aqui, também, é importante explicitar esta situação.

A compreensão de que as revisões sistemáticas são necessárias para fornecer experimentos controlados se refletiu em um rápido aumento no número de relatórios de revisões sistemáticas sendo publicados em papel e eletronicamente ([DARE](#); [The Cochrane Collaboration](#)). Estes são usados (i) para informar a prática clínica, por exemplo, através da BMJ publication [Clinical Evidence](#) e do [Scottish Intercollegiate Guidelines Network](#); (ii) para avaliar quais tratamentos médicos têm uma boa relação custo-benefício, por exemplo, pelo Instituto Nacional de Saúde e Excelência Clínica ([National Institute for Health and Clinical Excellence](#)) e (iii) para atender às necessidades dos pacientes quanto à informações confiáveis sobre os efeitos de tratamentos, por exemplo, através do [Informed Health Online](#) e da Biblioteca Nacional de Saúde ([National Library for Health](#)).

Negócios inacabados

Estes e outros desenvolvimentos semelhantes mostram que a importância das revisões sistemáticas foi aceita por aqueles que estão tentando melhorar o acesso à evidência necessária para informar as opções no atendimento à saúde. Entretanto, há ainda um longo caminho a seguir: Estima-se que a produção atual da Colaboração Cochrane de vários milhares de revisões sistemáticas terá de aumentar para mais de 10.000 para cobrir as evidências existentes (Mallett and Clarke 2002), e depois se manter atualizada à medida que novas evidências apareçam. Na verdade, o editor de uma revista sugeriu que deveria haver uma suspensão de todas as novas pesquisas até estarmos em dia com aquilo que as evidências existentes podem nos dizer (Bausell 1993).

Os responsáveis por fornecer fundos para as pesquisas devem assegurar-se de que os recursos sejam fornecidos para lidar com este acúmulo, e que novos estudos só sejam apoiados se as revisões sistemáticas das evidências existentes mostrarem que mais estudos são necessários, e que estes foram planejados para considerar as lições de pesquisas anteriores. Se os editores de revistas pretendem servir melhor os seus leitores, eles devem seguir o exemplo da The Lancet e assegurar que os relatórios de novos estudos deixem claro que contribuição a nova evidência trouxe para uma revisão sistemática atualizada de todas as evidências relevantes. (Young and Horton 2005).

Uma maior disponibilidade de revisões sistemáticas atualizadas está melhorando a qualidade das informações sobre os efeitos de tratamentos, mas as conclusões das revisões sistemáticas não devem ser aceitas

indiscriminadamente. Diferentes revisões, aparentemente, tratando da mesma questão sobre os tratamentos às vezes chegam à conclusões diferentes. Seus autores são humanos e devemos estar cientes de que eles podem selecionar, analisar e apresentar as evidências de maneiras que justifiquem seus preconceitos e interesses. A evolução contínua de métodos confiáveis para a preparação e a manutenção das revisões sistemáticas ajudarão a tratar deste problema, mas não podemos esperar que eles acabem de vez com o problema.

Embora o crescimento nos números de revisões sistemáticas tenha aumentado a disponibilidade dos principais experimentos controlados de tratamentos no atendimento à saúde, estas revisões, freqüentemente, também revelam a baixa qualidade e a irrelevância de muitas pesquisas sobre os efeitos de tratamentos. Conforme o comentário de um editorialista sobre "o escândalo da pesquisa médica de baixa qualidade", precisamos de menos pesquisas, pesquisas de melhor qualidade e pesquisas feitas pelos motivos certos (Altman 1994). É quase impossível que isto seja alcançado sem uma maior compreensão pública do fundamento lógico e das características dos experimentos controlados, e maior influência e envolvimento público em todas as fases dos experimentos controlados de tratamentos. A divulgação deste programa depende das incertezas sobre os efeitos de tratamentos sendo confrontadas por novas alianças de pacientes e médicos (Chalmers 2004; www.duets.nhs.uk; [James Lind Alliance](http://www.jameslindlibrary.org)).

O público e os profissionais da saúde serão mais bem servidos quando tiverem fácil acesso às revisões sistemáticas atualizadas de todas as evidências confiáveis e relevantes que tratam de importantes incertezas sobre os efeitos de tratamentos, e às informações sobre as pesquisas em andamento que tratam destas incertezas. (Smith and Chalmers 2001).

Cite as: Editorial commentary (2007). Avaliações sistemáticas atualizadas das evidências relevantes e confiáveis. The James Lind Library (www.jameslindlibrary.org).

Select essay:

Referências

Altman (1994). The scandal of poor medical research. *BMJ* 308: 283-284.

Bausell BB (1993). After the meta-analytic revolution. *Evaluation and the Health Professions* 16: 3-12.

Bunker JP, Frazier HS, Mosteller F (1994). Improving health: measuring effects of medical care. *Milbank Quarterly* 72: 225-258.

Chalmers I (2004). Well informed uncertainties about the effects of treatments: how should clinicians and patients respond? *BMJ* 328: 475-476.

Lind J (1753). A treatise of the scurvy. In three parts. Containing an inquiry into the nature, causes and cure, of that disease. Together with a critical and chronological view of what has been published on the subject. Edinburgh: Printed by Sands, Murray and Cochran for A Kincaid and A Donaldson.

Mallett S, Clarke M (2002). The typical Cochrane Review. *International Journal of Technology Assessment in Health Care* 18: 820-823.

Smith R, Chalmers I (2001). Britain's gift: a 'Medline' of synthesized evidence. *BMJ* 323: 1437-1438.

Young C, Horton R (2005). Putting clinical trials into context. *Lancet* 366: 107-8.

[Home](#)

[Conteúdo](#)

[Envie comentários](#)