

Viéses em estudos clínicos [18]

ANTÓNIO VAZ CARNEIRO

Centro de Estudos de Medicina Baseada na Evidência Faculdade de Medicina da Universidade de Lisboa, Lisboa, Portugal

Rev Port Cardiol 2011; 30 (02): 235-242

RESUMO

Os estudos clínicos servem fundamentalmente para obtenção de dados que suportem a prática, quer diagnóstica, quer terapêutica, quer prognóstica. Mas por melhor que o estudo tenha sido desenhado, implementado ou analisado nunca é possível garantir com toda a segurança a veracidade dos resultados. A explicação para esta imprecisão da estimativa pontual encontra-se em dois factores que alteram os cálculos: o erro aleatório e o erro sistemático (também designado por viés). O primeiro surge sem aviso e não é passível de modulação, enquanto o segundo – definido como uma introdução sistemática de erro num estudo – poderá ser removido (se não na totalidade, pelo menos parcialmente).

Neste artigo apresentamos o conjunto de viéses mais frequentes que se podem encontrar em estudos clínicos, quer nos que analisam uma intervenção terapêutica (os ensaios clínicos), quer em outros estudos experimentais, quer em estudos observacionais.

Palavras-chave

Erro em estudos clínicos;
Viéses em estudos clínicos; Validade interna;
Medicina baseada na evidência.

Bias in clinical studies

ABSTRACT

Clinical studies are designed fundamentally to obtain data to support practice, whether diagnostic, therapeutic or prognostic. No matter how well a study was designed, implemented or analyzed, it is never possible to guarantee the truth of its results with absolute certainty. This is due to two factors that can change the study's results: random error and systematic error (also called bias). The first appears without warning and cannot be modulated, while the second – defined as the systematic introduction of error into a study – can be removed (if not completely, at least partially).

In this paper we present the most common types of bias found in clinical studies, whether therapeutic (clinical trials), or with other experimental or observational designs.

Key words

Error in clinical studies; Bias in clinical studies;
Internal validity; Evidence-based medicine.

INTRODUÇÃO

Os estudos clínicos servem fundamentalmente para obtenção de dados que suportem a prática, por exemplo, quer no processo diagnóstico, quer na elaboração de um esquema terapêutico, quer no estabelecimento de um prognóstico.

A verdade por trás da resposta às questões clínicas permite obter uma estimativa o mais correcta possível do efeito da intervenção: por exemplo, qual é o benefício do uso dos inibidores da enzima de conversão (IECAs) nos doentes com insuficiência cardíaca por disfunção sistólica do ventrículo esquerdo?

Os estudos procuram determinar, com o máximo rigor possível, a verdade clínica subjacente. Mas de facto nunca será possível saber com absoluta precisão o impacto de uma terapêutica: por exemplo, vamos supor que a dimensão do efeito do tratamento num ensaio clínico (também designado como estimativa pontual) é apresentada como o de um risco relativo (RR) de morte de 0.80 (ou 80%). Neste caso, a verdade pode ter dois cenários: o primeiro, que o tratamento induz de facto uma redução do RR de 20%; ou, em alternativa, que esta redução pode ser superior ou inferior a 20%.

A explicação para esta imprecisão da estimativa pontual de efeito encontra-se em dois factores que podem alterar os resultados: o erro aleatório (também designado por chance) e o erro sistemático (também designado por viés)⁽¹⁾.

O erro aleatório surge sem aviso e não existe habitualmente possibilidade de o modular. O exemplo clássico é o da moeda perfeitamente equilibrada que, quando lançada ao ar 10 vezes, normalmente devolve um nº desigual de faces: por ex. 7 caras e 3 coroas (nós até ficaríamos surpreendidos se aparecessem 5 caras e 5 coroas...). E isto acontece apesar da probabilidade de cair caras ou coroas ser exactamente de 50%. De resto, um valor igual raramente se verifica: mesmo que aumentássemos o nº de lançamentos – digamos para 10 000 vezes – obteríamos com certeza por ex. 4980 caras *versus* 5020 coroas. Apenas no

INTRODUCTION

Clinical studies are designed fundamentally to obtain data to support practice, for instance in the process of diagnosis, in preparing a therapeutic regimen, or in establishing a prognosis.

The truth behind the answers to clinical questions enables the most correct possible estimate of the effect of an intervention: for example, what is the benefit of the use of angiotensin-converting enzyme inhibitors in patients with heart failure due to left ventricular systolic dysfunction?

Studies are designed to determine the underlying clinical truth with the greatest possible rigor. But in fact it is never possible to know the impact of a therapy with complete accuracy: for example the size of the treatment effect in a clinical trial (the point estimate) is presented as a relative risk (RR) of death of 0.80 (80%). In this case, two scenarios may reflect the reality: the treatment indeed reduces RR by 20%; or the reduction may be higher or lower than 20%.

This imprecision in the point estimate of the effect may be explained by two factors that can alter the results: random error (also known as chance), and systematic error (also known as bias)⁽¹⁾.

Random error appears without warning and cannot usually be modulated. The classic example is of a perfectly balanced coin that, when tossed 10 times, normally comes down heads or tails an unequal number of times, such as seven heads and three tails; it would in fact be surprising if it came down five times heads and five times tails, even though the probability of heads or tails is exactly 50%. Indeed, the numbers are rarely equal: if the number of tosses were increased to 10,000, we would certainly obtain, for example, 4980 heads and 5020 tails. Only with an infinite number of tosses would an exactly symmetrical result be found.

Random error is also found in clinical studies, whether diagnostic, therapeutic, prognostic or other, and so the study should be designed in such a way as to minimize its effects.

caso de um nº infinito de lançamentos se verificaria um resultado exactamente simétrico.

Esta hipótese de erro aleatório também se manifesta em estudos clínicos, sejam eles de diagnóstico, terapêutica, prognóstico, etc., pelo que o desenho do estudo deverá ter em conta a maior minimização possível daquele.

Pelo seu lado, o viés define-se como uma introdução sistemática de erro num estudo, podendo distorcer os seus resultados de maneira não aleatória⁽²⁾. Apesar de ser impossível eliminar todos os vieses dum estudo, é dever dos investigadores removerem o maior número possível daqueles. Como é óbvio, nem todos os vieses têm a mesma capacidade de invalidar os resultados de um estudo, mas essa análise deve ser feita com todo o cuidado, de maneira a identificar a real validade interna do estudo⁽³⁾.

Para detectar as fontes de vieses num estudo, é necessário ler com toda a atenção a sua secção do Material e Métodos e analisar os seguintes componentes⁽²⁾:

- Os métodos utilizados para as medições estão detalhadamente descritos, juntamente com um protocolo claro para a sua realização?
- Os observadores foram treinados para efectuar as medições, com descrição do método de treino a que foram sujeitos?
- As medições foram feitas discretamente e foram ocultadas em relação aos sujeitos do estudo (doentes), assim como aos investigadores e analisadores de resultados (ocultação simples, dupla ou tripla)?
- Realizaram-se medições emparelhadas (definição da fiabilidade teste-reteste) e a sua média calculada, garantindo-se a fiabilidade intra e inter-observador através de medições múltiplas?
- As medições foram comparadas com o "padrão-ouro" (a medição considerada verdadeira) e a sua validade assegurada por dados da literatura?
- A explicitação dos critérios de inclusão e exclusão no estudo é clara?
- A identificação dos *dropouts*, assim como das suas causas, e as análises de sensibilidade com inclusão/exclusão destes sujeitos foram efectuadas?

On the other hand, bias is defined as the systematic introduction of error into a study that can distort the results in a non-random way⁽²⁾. Although it is almost impossible to eliminate all sources of bias from a study, it is the job of the researcher to attempt to remove as much bias as possible. Obviously, not all bias will invalidate a study's results, but this analysis should be performed as carefully as possible in order to determine the study's real internal validity⁽³⁾.

To detect sources of bias in a study, its methods section should be read carefully and the following components analyzed⁽²⁾:

- Were the methods for making all the measurements fully described with a clearly defined protocol for making these measurements?
 - Were the observers trained to make the measurements and was this training described?
 - Were measurements made unobtrusively, were the subjects blinded to the measurement being made, and were the observers and interpreters blinded to the results (single, double or triple blinding)?
 - Were paired measurements made (test-retest reliability) or averaged, and was intra- or inter-observer reliability of repeated measurements measured?
 - Were the measurements checked against a gold standard (accepted as the true measurement) and checked for their validity through data from the literature?
 - Were the reasons for inclusion and exclusion spelled out?
 - Were dropouts and their causes identified, and was the effect on the study's sensitivity analyzed with them included and excluded?
 - Was the selection of primary and secondary outcomes explained and justified?
- In practical terms, a study should have good internal validity (i.e. it should measure what it was designed to measure); only if it does will its results be credible. But it should also have good external validity – the results should be generalizable to the general population; only if they are will they be useful⁽¹⁾.

Internal and external validity to some extent balance each other: when one increases, the other tends to decrease. A typical

• Está explicitada a justificação da selecção de indicadores (*outcomes*) primários e secundários?

Em termos práticos, um estudo deve possuir uma boa validade interna, isto é, medir o que é suposto medir (só assim os seus resultados poderão ser credíveis). Mas deve também possuir uma boa validade externa, isto é, os resultados devem poder ser generalizáveis à população em geral. Só assim os resultados poderão ser úteis⁽¹⁾.

As validades interna e externa estão como que em equilíbrio: quando uma aumenta, a outra tem tendência a diminuir. O exemplo típico é o do ensaio clínico, em que uma validade interna elevada - conseguida através de critérios de inclusão e exclusão muito restritos produzindo uma amostra muito homogénea - pode atingir a generalização dos resultados para doentes da prática diária, que são mais velhos ou mais novos, apresentam doenças concomitantes que interferem com a resposta ao tratamentos, têm pesos corporais diferentes, etc.⁽⁴⁾.

Neste artigo apresentamos o conjunto de viéses mais frequentes que se podem encontrar em estudos clínicos, quer nos que analisam uma intervenção terapêutica (os ensaios clínicos), quer em outros estudos experimentais quer em estudos observacionais.

Alguns tipos de viéses encontrados em estudos clínicos⁽¹⁾

Viés de amostragem (*sampling bias*): erro sistemático de estudo de uma amostra não-aleatorizada da população.

Viés de associado (*membership bias*): acontece quando o nível de saúde basal dos doentes integrados no estudo é diferente dos da população em geral, por pertencerem por exemplo, a ginásios ou equipas desportivas.

Viés de ausência de resposta (*non-response bias*): é o que se pode passar num estudo epidemiológico, por exemplo em que os doentes que não respondem aos inquéritos são normalmente diferentes em muitas características clínicas e sociais respondedores.

Viés de averiguação (*ascertainment bias*): é a falta sistemática de inclusão de todas as po-

example is a clinical trial with high internal validity resulting from strict inclusion and exclusion criteria that produce a homogeneous sample, which may not be generalizable to patients in daily clinical practice, who are older or younger, have concomitant pathologies that affect their response to treatment, have different body weight, or other differences⁽⁴⁾.

In this paper we present the most common types of bias found in clinical studies, whether therapeutic (clinical trials), or with other experimental or observational designs.

Some types of bias encountered in clinical studies¹

Sampling bias: systematic error in the study of a non-randomized sample of the population.

Membership bias: this occurs because the health of some group members differs in a systematic way from the general population, for example because they belong to health clubs or sports teams.

Non-respondent bias: this may be observed in an epidemiological study in which subjects who do not respond to a survey may be different in some fundamental way from those who respond.

Ascertainment bias: this is the systematic failure to include all potential supposedly representative classes or subgroups of patients when selecting a study sample, such as selecting hospitalized patients when those in primary care are also important.

Berkson's bias: is a specific bias that occurs when patients in a selected ward of a hospital are the control subjects. Their baseline characteristics may be so different from the cases that valid comparisons are impossible.

Classification bias: this occurs when study subjects are wrongly classified, such as when patients are diagnosed as having a disease when they do not in fact have it (due to error in the process of diagnosis).

Confounding bias: this is the erroneous attribution of causality to one of various risk factors associated with the disease found in the individual patient. For example, if exer-

tenciais classes ou subgrupos de pacientes supostamente representativos na constituição de uma amostra; por exemplo, seleccionar a população para estudo a partir de doentes hospitalares, quando os de cuidados primários também são importantes.

Viés de Berkerson (*Berkerson's bias*): é um viés específico que pode aparecer quando um doente internado numa enfermaria hospitalar é seleccionado para um estudo. As suas características basais podem ser tão diferentes, que invalidam as comparações com os doentes do grupo experimental.

Viés de classificação (*classification bias*): ocorre quando os sujeitos do estudo são erradamente classificados: por exemplo, quando é atribuído um diagnóstico de uma doença a um paciente que ele de facto não tem (por erro no processo diagnóstico).

Viés de confundimento (*confounding bias*): é a errada atribuição de causalidade a um de diversos factores de risco associados à doença que se encontrem presentes no doente individual. Por exemplo, a atribuição do benefício do exercício físico na diminuição da incidência de doença cardiovascular deverá garantir que todos os outros factores de risco potenciais estejam controlados (por exemplo, que entre os dois grupos estudados a taxa de tabagismo seja idêntica).

Viés de detecção (*detection bias*): é uma forma de viés de selecção em que os doentes são preferencialmente incluídos no estudo por terem sido expostos a um factor de risco determinado. Este factor de risco induz a pesquisa da doença e é considerado causal em relação a esta.

Viés de espectro (*spectrum bias*): acontece quando só os doentes com sintomas clássicos ou severos são incluídos no estudo. Por exemplo, doentes com AVC hemorrágico e com sintomas clínicos típicos de cefaleia de alto risco apresentam consistentemente uma TC subsequente que é positiva. Outros doentes com cefaleias de baixo risco não têm TC positivas.

Viéses de ensaio clínico: 1) erro sistemático nas medições dos dados; 2) erro sistemático nos cálculos estatísticos (médias, taxas, medidas de associação, etc.); 3) fraquezas metodológicas do estudo (na recolha,

em que se considera que tem um efeito benéfico sobre a incidência de doença cardiovascular, é necessário assegurar que todos os outros factores de risco potenciais são controlados (por exemplo, assegurando que a taxa de fumo nos dois grupos é a mesma).

Detection bias: this is a form of selection bias that preferentially includes patients in a study if they have been exposed to a particular risk factor. This risk factor precipitates a search for the disease and is then blamed for causing the disease.

Spectrum bias: this occurs when only patients with classical or severe symptoms are selected for a study. For example, patients with hemorrhagic stroke and typical symptoms of high-risk headache will almost all have a positive CT. Patients with low-risk headache are much less likely to have a positive CT.

Types of bias in clinical trials: 1) systematic error in the measurement of data; 2) systematic error in statistical calculations (means, rates, measures of association, etc.); 3) methodological weaknesses (in collection, analysis, interpretation, publication or review of the data); 4) wrong analytical techniques applied to the constituent elements of a clinical study or trial; 5) the preconceptions of the researchers; 6) contamination (both groups receive the experimental therapy, or receive different basic treatment according to their baseline risk); 7) losses to follow-up during the study, when information on these subjects could be highly relevant to the final analysis; and 8) errors in comparing the results in the two groups caused by different dropout rates or exclusions, such as those caused by the side-effects of the experimental therapy (attrition bias).

Interpretation bias: this is caused by erroneous or incomplete inferences and speculation, when not all possible interpretations consistent with the facts are considered, or when exceptions are ignored.

Recall bias: this is the lack of information – especially important in retrospective or case-control studies – due to a systematic error in individuals' ability when questioned

análise, interpretação, publicação ou revisão dos dados); 4) erradas técnicas analíticas aplicáveis aos factores constitutivos de um ensaio/estudo clínico; e 5) desvios provocados por preconceitos dos investigadores; 6) ambos os grupos recebem a terapêutica do grupo experimental ou recebem tratamento de base diferente segundo o seu risco basal (contaminação); 7) perda de seguimento de doentes durante a execução do estudo (a informação destes doentes pode ser muito importante para a análise final); 8) erro na comparação de resultados de doentes nos dois grupos em comparação por diferenças de *dropouts*, ou exclusão daqueles – por exemplo devido a efeitos secundários da droga experimental (*Viés de atrito - attrition bias*).

Viés de interpretação (*bias of interpretation*): erro proveniente de inferências e especulações erradas ou incompletas (não considerar todas as possíveis interpretações consistentes para os factos, ou ignorar os casos de excepção).

Viés de lembrança (*recall bias*): é a falta de informação suficiente – especialmente importante em estudos retrospectivos ou caso-controlo - por dificuldades dos sujeitos, quando questionados, poderem recordar com precisão os factos relevantes ocorridos no passado. Isto acontece tipicamente no caso de um efeito secundários de um medicamento, em que um doente afectado tem tendência a recordar com mais precisão esse evento do que um paciente que nunca experimentou episódio semelhante.

Viés de observador (*observer bias*): é a distorção (consciente ou não) da percepção da comunicação das medidas pelo observador. O caso típico é o efeito *Hawthorne*, em que os sujeitos sob observação alteram o seu comportamento quando sabem estar a ser analisados. Outro exemplo é o de dois médicos que tratam os doentes de maneira diferente conforme estão convencidos que estes estão a fazer placebo (aumentando a intensidade terapêutica) ou substância activa (diminuindo-o).

Viés de participante (*subject bias*): é uma distorção constante da medida pelo indivíduo estudado. Por exemplo, certos sujeitos da investigação procuram dizer aos investigadores

to recall relevant facts accurately. A typical example is a subject who has suffered a secondary effect of a drug, who is more likely to remember the event than one who has never had a similar experience.

Observer bias: this is the conscious or unconscious distortion in perception of reporting the measurement by an observer. A classic example is the Hawthorne effect, in which being observed during the process of making measurements changes the behavior of the subject. Another example is that of two physicians who treat subjects differently according to whether they believe they are taking the placebo (and therefore give them more intensive treatment) or the active substance (and therefore give them less intensive treatment).

Subject bias: this is a constant distortion of the measurements by the subject. For example, subjects will tell researchers what they think they want to hear, or change their behavior during the study according to the instructions they have been given.

Publication bias: this arises from the tendency of medical journals to publish studies that show a positive result (especially if it is considered “news”), compared to studies that show no effect or a negative result (especially if this confirms known findings). This bias can diminish the perception of an association between two factors, such as a risk factor and a pathology, or of the efficacy of a new drug, which may appear more effective than it really is. A second important consequence is that publication bias can cause problems in meta-analyses, by making it impossible to obtain an overall idea of the effect (or lack of effect) of the intervention.

Referral bias: this arises when studies that include patients in tertiary referral centers are extrapolated to patients treated in primary health care. The diagnostic characteristics of most tests were defined in hospitalized populations but are then applied freely in the community; since the prevalence of disease in the latter is by definition lower, this lowers the test’s discriminatory power.

Selection bias: patients are selected in a manner that will systematically influence the

aquilo que eles querem ouvir, ou então alteraram o seu comportamento sob estudo em conformidade com as instruções recebidas.

Viés de publicação (*publication bias*): consiste no facto dos editores de revistas médicas terem uma tendência a publicar mais frequentemente estudos que apresentam resultados «positivos» (especialmente se estes forem considerados «novidades»), em oposição aos estudos com resultados «negativos», não significativos (especialmente se confirmam dados já conhecidos na literatura). Este viés pode diminuir a percepção da existência de uma associação entre dois factores (por exemplo um factor de risco com uma patologia) ou da eficácia terapêutica de uma molécula nova (que parece mais eficaz do que na realidade é). Uma segunda consequência importante é de ser uma fonte de problemas nas meta-análises, já que não se poderá ficar com uma ideia do efeito (ou não-efeito) global da intervenção.

Viés de referenciação (*referral bias*): acontece em estudos que incluem amostras de doentes de centros terciários e extrapolam para doentes de centros primários de saúde. Exemplo: as características diagnósticas da maior parte dos testes conhecidos foram definidas em populações hospitalares, sendo aplicadas livremente em contextos comunitários. Como a prevalência de doença nestes últimos contextos é, por definição, mais baixa, a capacidade discriminativa dos testes é inferior.

Viés de selecção (*selection bias*): os doentes são seleccionados de maneira que influenciam sistematicamente os resultados do estudo, por exemplo devido à existência de diferenças sistemáticas nas características dos sujeitos seleccionados versus os não seleccionados (os doentes voluntários *versus* os pagos podem apresentar diferenças substantivas em termos da doença a estudar, por ex.).

Viés de suspeita de exposição (*exposure suspicion bias*): é um viés de lembrança induzido pelo próprio investigador que, quando questiona o doente acerca da exposição, formula a questão em termos de favorecer o viés de lembrança.

outcome of the study, due to the existence of systematic differences in the characteristics of those selected compared to those not selected, such as subjects who are volunteers or paid to be in the study may have different characteristics in terms of the disease in question.

Exposure suspicion bias: this is a type of recall bias that occurs on the part of the researcher. When asking subject patients about exposure, researchers might phrase the question in ways that encourage recall bias.

Volunteer bias: this occurs because patients or controls who volunteer to participate in clinical trials usually have different characteristics from those selected randomly. There is evidence that those who volunteer for studies on preventive measures have a healthier lifestyle than those selected randomly from a non-volunteer database, and will therefore tend to have better outcomes.

¹ The interested reader can consult standard works on the subject^(5, 6, 2, 7-9).

Viés de voluntariado (*volunteer bias*): este viés aparece porque normalmente os doentes (ou controlos) que se oferecem como voluntários para participar num ensaio clínico possuem características diferentes dos controlos seleccionados ao acaso; por exemplo, existe evidência que pacientes que se oferecem para estudos sobre medidas preventivas poderem ter, à partida, um estilo de vida mais saudável que pacientes seleccionados ao acaso de uma base de dados não voluntária, pelo que os resultados naquele grupo serão sempre melhores.

¹ Os leitores interessados neste assunto poderão consultar algumas fontes clássicas ^(5,6,2,7,8,9)

Pedido de Separatas
Address for Reprints

António Vaz Carneiro
Centro de Estudos de Medicina
Baseada na Evidência
Faculdade de Medicina de Lisboa
1649-035 Lisboa
e-mail: ave@fm.ul.pt

BIBLIOGRAFIA / REFERENCES

1. Esuyett et al. User's Guides to the Medical Literature. 2nd. New York: McGraw-Hill; 2008
2. Mayer D. Essential evidence-based medicine. 1st. Cambridge: Cambridge University Press; 2004
3. Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet* 2002;359:248-252.
4. Rottiwel PM, Treating individuals. From randomised trials to personalised medicine. 1st. Edinburgh: Elsevier; 2007
5. Carneiro AV. Glossário de termos e conceitos utilizados pela Cardiologia baseada na evidência científica. Parte I - Diagnóstico, terapêutica, iatrogenia, generalidades. *Rev Port Cardiol* 2001;20:99-103.
6. Carneiro AV. Glossário de termos e conceitos utilizados pela cardiologia baseada na evidência científica. Parte II : Ensaio clínicos. *Rev Port Cardiol* 2001;20:203-210.
7. Katz DL. Clinical epidemiology & evidence-based medicine. 1st. Thousand Oaks: Sage Publications; 2001
8. Greenhalgh T. How to read a paper. London: BMJ Books; 2010
9. Elwood M. Critical appraisal of epidemiological studies and clinical trials. 3d. Oxford: Oxford University Press; 2007