
Conferência / Conference

**COMO AVALIAR A INVESTIGAÇÃO CLÍNICA.
O EXEMPLO DA AVALIAÇÃO CRÍTICA DE UM ENSAIO CLÍNICO**

ANTÓNIO VAZ CARNEIRO

Resumo

O papel da ciência na prática clínica é, hoje em dia, absolutamente insubstituível. A publicação permanente de estudos e ensaios clínicos produz evidência (prova científica) de boa qualidade, sobre a qual é possível o médico tomar decisões sólidas, mesmo que num contexto de incerteza e risco. Para além disso, a combinação entre a gestão de recursos cada vez mais escassos e dispendiosos por um lado, com responsabilização dos médicos por parte da sociedade na prestação de cuidados eficazes mas custo-efectivos, por outro, cria novas exigências de rigor e racionalização da prática médica. Hoje em dia, a evidência científica proveniente de estudos clínicos pode caracterizar-se – em relação à existente digamos há 30 anos atrás – em três pontos: 1) existe muito maior quantidade de estudos; 2) a qualidade destes é muito melhor; e 3) estes podem ser rápida e eficazmente localizados (através de meios informáticos que pesquisam as grandes bases de dados biomédicas). Portanto a questão não deverá ser “*Existe uma resposta para a minha questão?*” (a resposta será na maior parte dos casos afirmativa), mas sim “*Estes dados que encontrei são fiáveis?*”. Por estas (e outras) razões, torna-se absolutamente fundamental que o médico saiba analisar a qualidade metodológica dos artigos que seleccionou para ler.

Neste artigo apresentamos, como exemplo de avaliação da investigação clínica, um esquema de análise crítica de ensaios clínicos.

Summary

The role of science in medical practice is nowadays absolutely irreplaceable. The continuous publication of studies and clinical trials produces evidence of good quality, based on which physicians can base their decisions, even in an uncertain and risky environment. Moreover, the combination of rare and expensive resource management with professional accountability of doctors in efficacious and cost-effective care creates new rational needs and demands. Nowadays, scientific research from clinical studies can be looked upon from three standpoints of view (as opposed from 30 years ago): 1) there is a much bigger number of studies; 2) their quality is higher and 3) they can be located easily (through search of electronic databases). So the question is not “*Is there an answer to my questions?*” (most of the time the answer will be yes), but “*Are these data I locate reliable?*”. For these (and other) reasons, it is absolutely mandatory that doctors know how to critically appraise the clinical evidence from studies. In this paper we present a scheme of critical appraisal of clinical trials, as an example of evaluation of clinical research.

GE - J Port Gastroenterol 2008; 15: 30-36

INTRODUÇÃO

A prática clínica no início do terceiro milénio é complexa e incerta, e o médico prático é posto em confronto com problemas de conhecimentos diariamente. Os novos avanços diagnósticos e terapêuticos em medicina processam-se a um ritmo acelerado, criando problemas de actualização e aplicação prática a quem tem a responsabilidade da assistência médica a doentes internados em hospitais, observados em consultas ou avaliados em serviços de urgência⁽¹⁾. O papel da ciência na prática clínica é, hoje em dia, absolutamente insubstituível. A publicação permanente de estudos e ensaios clínicos produz evidência (prova científica) de boa qualidade, sobre a qual é possível o médico

tomar decisões sólidas, mesmo que num contexto de incerteza e risco⁽¹⁾.

Para além disso, a combinação entre a gestão de recursos cada vez mais escassos e dispendiosos por um lado, com responsabilização dos médicos por parte da sociedade na prestação de cuidados eficazes mas custo-efectivos, por outro, cria novas exigências de rigor e racionalização da prática médica⁽²⁾.

Na base da resolução dos problemas acima apontados encontra-se a necessidade de obtenção e síntese de informação clínica válida e relevante que sirva de base à actividade do médico que procura resolver os problemas clínicos que se lhe colocam diariamente. A questão essencial é então a de saber como podem os médicos apren-

der as inovações e dominar a informação de modo a introduzir (eventuais) mudanças na sua prática que, em última análise, irão beneficiar os seus doentes.

A investigação clínica define-se como sendo aquela que é realizada em doentes (ou pacientes), com intervenções que são relevantes para aqueles, em condições que possam ser realizadas no dia-a-dia e cujos resultados se destinam a servir de base à decisão de diagnosticar, tratar ou prognosticar as patologias humanas.

As questões que a investigação procura responder são instrumentais na base do desenho dos estudos que as irão tentar esclarecer: para questões sobre a eficácia de uma intervenção terapêutica, o desenho ideal é o ensaio clínico aleatorizado, prospectivo e controlado; mas se procurarmos definir as características diagnósticas de um novo teste (em comparação com o *gold standard* para aquela patologia) então deve-se utilizar um estudo seccional cruzado; para o prognóstico, um estudo prospectivo de coorte, para a etiologia, um estudo caso-controlo, e assim por diante.

A avaliação crítica da evidência científica consiste em analisar e classificar os estudos em termos da sua validade, da sua importância e da sua aplicabilidade clínica, para a inclusão (ou exclusão) da informação assim obtida como complemento na decisão final sobre o doente (dependendo do grau de confiança que se deposita na estrutura metodológica dos estudos). A necessidade da utilização de um critério de avaliação rigoroso advém do facto da maior parte da investigação publicada (de grandes dimensões) possuir pouco rigor metodológico e por vezes ser também pouco relevante em termos da sua utilidade prática.

A avaliação crítica da evidência consegue-se através das respostas a um conjunto de questões e regras predefinidas. Embora lógicas, nem todas estas são auto-explanatórias, pelo que necessitam de instruções detalhadas para uma aplicação correcta a estudos terapêuticos ou de prevenção⁽³⁻⁵⁾, diagnósticos⁽⁶⁻⁸⁾, de iatrogenia⁽⁹⁾, de estabelecimento de prognóstico⁽¹⁰⁾ ou de revisão⁽¹¹⁾, para apenas citar alguns tipos. Por exemplo, uma questão metodológica fundamental na validade de um ensaio clínico é a de saber se os doentes nos grupos experimental ou de controlo foram nestes incluídos através de um processo de aleatorização devidamente implementado. Mas se olharmos para um estudo diagnóstico, então o que é importante é ver se todos os doentes fizeram ambos os testes em comparação. E num estudo prognóstico, há que analisar se o grupo (coorte) que se pretende seguir no tempo possui doentes com todas as gamas de gravidade de doença.

E estas são apenas questões sobre a validade. Se pensarmos na importância e aplicabilidade dos resultados, então as questões serão outras. Por exemplo, se uma intervenção terapêutica reduz a mortalidade de 12,3% para 11,2%

e esta diferença for *estatisticamente* significativa, será *clínicamente* significativa isto é, será suficientemente importante para ser recomendada? E se um teste diagnóstico for tão complexo de aplicar que só possa ser feito em hospitais terciários, será possível disseminá-lo em todo o sistema de saúde (aplicabilidade)?

Poder-se-á perguntar porque é que necessitamos de utilizar um critério de avaliação tão rigoroso, em vez de nos fiarmos na nossa intuição. A resposta advém do facto da maior parte da investigação publicada (de grandes dimensões) possuir pouco rigor metodológico e por vezes ser também pouco relevante em termos da sua utilidade prática. Daí a necessidade de um processo de avaliação crítica muito rigoroso⁽¹²⁾.

Neste artigo iremos dar o exemplo da avaliação de um ensaio clínico (*randomized controlled trial* – RCT), devido à importância deste tipo de estudo na literatura biomédica.

AVALIAÇÃO CRÍTICA DE UM ENSAIO CLÍNICO

As questões (há quem lhe chame “guias”) a que um ensaio clínico deverá responder para ser correctamente avaliado estão descritas no Quadro I. Estas são questões metodológicas, que procuram analisar a validade do estudo (isto é, se o seu desenho permite responder à questão colocada), a importância dos resultados (isto é, se são clinicamente – para além de estatisticamente – significativos) e a aplicabilidade dos achados a outros doentes que naturalmente não participaram no estudo⁽¹³⁾.

Em termos práticos, cada questão é classificada em 4 hipóteses:

- resposta afirmativa = sim (codificada com 2)
- pouco claro/possivelmente (codificada com 1)
- resposta negativa = não (codificada com 0)
- não aplicável

No final de responder às 20 questões, obtém-se a soma de todas as questões individuais e calcula-se um *score* final, segundo o esquema:

- *Score* total (soma dos *scores* atribuídos) _____ [A]
- N^o de questões aplicáveis (máx. 20) _____ [B]
- *Score* máximo possível (2 x B) _____ [C]

Os ensaios clínicos classificados com uma pontuação de 75% ou mais são considerados de boa qualidade. Todos os outros são - em princípio - excluídos, já que não se revelam como tendo qualidade suficiente para serem considerados para a prática clínica.

O restante deste artigo irá descrever e justificar (de maneira sucinta) cada uma das questões a que um ensaio clínico deverá responder na sua avaliação crítica, para poder ser considerado válido ou não.

Quadro I - Grelha para avaliação crítica de um artigo descrevendo um ensaio clínico.

VALIDADE DOS RESULTADOS	S	?	N	n/a
1. a gama de doentes foi bem definida?	2	1	0	n/a
2. os critérios de inclusão e exclusão são lógicos e claros?	2	1	0	n/a
3. o diagnóstico da doença estava bem caracterizado?	2	1	0	n/a
4. os doentes foram aleatorizados?	2	1	0	n/a
5. o método de aleatorização foi explicado?	2	1	0	n/a
6. a distribuição foi ocultada?	2	1	0	n/a
7. os doentes foram analisados nos grupos para os quais tinham sido aleatorizados inicialmente (intenção-de-tratar)?	2	1	0	n/a
8. a dimensão da amostra foi estatisticamente calculada?	2	1	0	n/a
9. os doentes nos grupos em comparação eram semelhantes em termos dos seus factores de prognóstico conhecidos?	2	1	0	n/a
10. com excepção do tratamento em estudo, todos os doentes foram tratados da mesma maneira?	2	1	0	n/a
11. foi ocultado aos doentes o grupo a que pertenciam?	2	1	0	n/a
12. foram ocultados aos investigadores os grupos em estudo?	2	1	0	n/a
13. foram ocultados aos analisadores dos dados os grupos em estudo?	2	1	0	n/a
14. o seguimento (<i>follow-up</i>) final foi superior a 80%?	2	1	0	n/a
IMPORTÂNCIA DOS RESULTADOS				
15. a dimensão do efeito terapêutico (RRR, RRA, NNT) foi importante?	2	1	0	n/a
16. a estimativa do efeito é suficientemente precisa (IC)?	2	1	0	n/a
17. esse efeito tem importância clínica?	2	1	0	n/a
APLICABILIDADE DOS RESULTADOS				
18. os doentes do estudo são semelhantes aos da prática clínica do médico individual?	2	1	0	n/a
19. foram considerados todos os resultados clínicos importantes?	2	1	0	n/a
20. os benefícios do tratamento sobrepõem-se aos potenciais riscos e custos da sua implementação?	2	1	0	n/a

Validade dos Resultados

A gama de doentes foi bem definida?

A importância da determinação cuidadosa da gama de doentes de um RCT constitui um passo absolutamente crucial para análise da respectiva validade. Especialmente no que se refere ao seu risco de base (i.e. gravidade da doença), esta análise permite definir que representatividade é que esta amostra possui em relação à população em geral. Por exemplo: num ensaio clínico de um novo medicamento para a hepatite C, será necessário garantir que a amostra possui doentes com um quadro assintomático, mas também com insuficiência hepática (e todos os estádios intermédios). Só deste modo poderemos analisar o impacto da nova droga no prognóstico destes doentes.

É claro que se pode seleccionar um grupo de doentes muito homogéneo – por ex. infectados com HCV e com determinadas alterações morfológicas na biopsia hepática – e estudar apenas esses. A utilidade da informação sobre este subgrupo específico contrastará com a menor generalização dos resultados a todos os infectados com hepatite C ⁽⁵⁾.

Os critérios de inclusão e exclusão são lógicos e claros?

Um RCT destina-se a estudar uma intervenção numa população determinada. Como não é habitualmente possível estudar a totalidade da população (devido à sua di-

mensão), então é necessário definir quais serão os doentes que irão constituir a amostra. Isso consegue-se através da definição de critérios de inclusão e exclusão.

Os critérios de inclusão incluem tipicamente o grupo etário dos doentes, o seu sexo, o diagnóstico clínico e comorbilidades (doenças concomitantes). Os critérios de exclusão definem características dos doentes que precludem, por questões de segurança, a sua inclusão no estudo (por exemplo existência de insuficiência renal, idade superior a 75 anos, alergias conhecidas, etc.) ⁽¹⁴⁾.

O diagnóstico da doença estava bem caracterizado?

A definição precisa do diagnóstico é de central importância nos ensaios clínicos. Com efeito, para podermos aplicar aos nossos doentes os achados dum RCT (a sua validade externa), é absolutamente necessário que o diagnóstico definido no estudo seja claro e não ambíguo, já que pode atingir o seu poder estatístico ⁽¹⁵⁾.

Nalgumas situações, os critérios de inclusão e exclusão quase definem o diagnóstico, especialmente se forem muito específicos, mas seja como for, se existir alguma a incerteza sobre o diagnóstico, esta deverá sempre ser claramente indicada.

Os doentes foram aleatorizados?

Num ensaio clínico válido, os doentes devem ser distribuídos pelos grupos de análise (experimental/ais, controlo) baseados num processo aleatório imprevisível. Por

outras palavras, na distribuição aleatória cada participante tem uma probabilidade conhecida de receber um dos tratamentos em estudo antes de ser incluído no grupo respectivo, mas o tratamento que vai receber é determinado pelo acaso e não deve ser previsto antes ⁽¹⁶⁾.

A principal vantagem da aleatorização é a distribuição dos factores de confundimento desconhecidos: por exemplo, a genética parece desempenhar um papel importante na incidência de doença inflamatória do intestino (sabe-se que os nórdicos têm uma incidência aumentada da doença). Num estudo multicêntrico sobre uma nova terapêutica para a colite ulcerosa ou doença de Crohn, é fundamental que o risco genético de base seja o mesmo entre os grupos em estudo, pelo que só a aleatorização é que pode garantir (mesmo que não em 100% dos casos) um número análogo de doentes de ascendência nórdica entre os grupos.

O método de aleatorização foi explicado?

O método de aleatorização clássico elege o doente individual como unidade e procura distribuir cada um deles de per si, isto é, cada novo doente que é recrutado tem exactamente as mesmas hipóteses de ser incluído num dos grupos em estudo. Existem no entanto técnicas de aleatorização designadas como *restringidas*, que se destinam a manter um determinado equilíbrio entre os grupos, em termos de dimensão ou características específicas. Designa-se por blocos, estratificada e minimizada. Qualquer que seja o seu tipo, o método aleatorização de deve sempre ser descrito ⁽¹⁷⁾.

A distribuição foi ocultada?

Após uma aleatorização bem sucedida torna-se necessário garantir que a distribuição dos doentes pelos grupos para que foram inicialmente aleatorizados é feita sem erros. O ideal é que a distribuição seja feita num centro distante do qual se processa o estudo (contactando o investigador por exemplo pelo telefone quando quiser saber que tratamento irá aplicar) ⁽¹⁸⁾. Isto é necessário para garantir que a aleatorização é respeitada.

Os doentes foram analisados nos grupos para os quais tinham sido aleatorizados inicialmente (intenção-de-tratar)?

A abordagem intuitiva em relação aos doentes que não seguiram as instruções de tratamento que lhe foram indicadas, é de os retirar da análise final. Parecendo certa, esta decisão é fundamentalmente errada, já que subverte a vantagem da aleatorização, impede uma visão mais realista do impacto da nova droga e permite a introdução de viéses no estudo. Assim, independentemente de qual foi

a intervenção efectivamente cada doente recebeu, para a análise final ele deverá ser incluído no grupo para que inicialmente foi aleatorizado ⁽¹⁹⁾.

A dimensão da amostra foi estatisticamente calculada?

A dimensão de uma amostra de um ensaio clínico não pode ser definida de qualquer maneira, na base de convicções e crenças dos investigadores. Por razões clínicas e estatísticas, a dimensão amostral deve ser calculada com rigor, para que o estudo possa provar o que se propôs e, se o resultado final for negativo, não ter sido por um problema de a amostra insuficiente.

Os factores que entram em conta no cálculo da amostra são: 1) a percentagem de *outcomes* esperado em cada grupo; 2) o erro tipo I (nível α); 3) o erro tipo II (nível β); e, para *outcomes* contínuos, 4) o desvio padrão.

Os doentes nos grupos em comparação eram semelhantes em termos dos seus factores de prognóstico conhecidos?

Em todos os artigos reportando ensaios clínicos deverá existir um quadro que descreve, lado a lado, as características demográficas, clínicas, laboratoriais, etc. dos grupos em estudo. Idealmente aquelas serão sobreponíveis nos dois (ou mais) grupos, isto é, a média de idades será a mesma, a percentagem de homens idem, as características da doença também, etc. Isto é necessário para podermos assacar as eventuais diferenças nos benefícios e riscos do tratamento unicamente à intervenção (e não a um prognóstico de base diverso) ⁽²³⁾.

É claro que, mesmo com uma aleatorização competentemente desenhada e implementada, pode acontecer que – por pura má sorte – os grupos não estejam equilibrados. Neste caso, dever-se-á analisar se as diferenças são suficientemente importantes para anular os resultados ou se, pelo contrário, não influenciam a alocação de causalidade dos resultados.

Com excepção do tratamento em estudo, todos os doentes foram tratados da mesma maneira?

Um dos aspectos essenciais na análise da eficácia de uma intervenção é que ela possa ser inequivocamente responsabilizada pelas alterações medidas. Se por acaso o tratamento concomitante que os doentes sempre têm que fazer diferir entre os grupos, então torna-se muito difícil identificar a causa do benefício demonstrado: será pela nova droga, ou pelas alterações instituídas na terapêutica de base? É esta a razão porque há que garantir que os tratamentos não experimentais são análogos nos grupos em estudo.

Foi ocultado aos doentes o grupo a que pertenciam?

A razão porque é necessário que os doentes de um ensaio clínico não saibam o que estão a tomar (placebo ou substância activa) é porque essa informação pode alterar a sua percepção de efeito da intervenção. Se, por exemplo, estamos a testar um medicamento para o colesterol, essa informação poderá não ser muito relevante para o doente (afinal estamos a medir uma substância no soro que, em princípio, não será influenciada pelo conhecimento do que o doente está a tomar – o chamado *hard endpoint*). Mas se estivermos a estudar um novo medicamento para a depressão, por ex., então se o doente souber que está a fazer uma substância activa poderá referir superiores melhorias do que se lhe for dito que está a fazer um placebo - uma substância inerte (a melhoria da depressão, medida por uma escala validada, constitui um chamado *soft endpoint*).

Um estudo em que apenas os doentes não sabem o que estão a tomar designa-se por “com ocultação simples”⁽²⁴⁾. A ocultação destina-se a evitar os erros sistemáticos (viéses) de um estudo.

Foram ocultados aos investigadores os grupos em estudo?

Os investigadores, por seu turno, devem ser ocultados aos tratamentos que administram para evitar o viés de performance. Isto é fundamental, porque muitas vezes o conhecimento prévio de qual tratamento (ou não) o doente irá receber altera o comportamento do investigador responsável por implementar o estudo⁽¹⁴⁾. Por ex: se aquele souber que o doente irá receber um placebo, poderá ter a tentação de aumentar a terapêutica concomitante que o doente está a fazer, deste modo alterando os resultados do ensaio. Um estudo em que os doentes e os investigadores não sabem o que estão a tomar designa-se por “com dupla ocultação”⁽²⁴⁾.

Foram ocultados aos analisadores dos dados os grupos em estudo?

Se a interpretação sobre a evolução dos *outcomes* que estiverem a ser medidos for algo subjectiva, então os analisadores de resultados não poderão saber quem fez o placebo ou a substância activa. Existem estudos que demonstram que cardiologistas encarregados de analisar a evolução da função ventricular esquerda de doentes com insuficiência cardíaca de um ensaio clínico, conhecedores da terapêutica de cada doente, classificaram consistentemente como mais melhorada a função ventricular esquerda dos doentes a fazer terapêutica activa... Quando ocultados para essa informação, a sua classificação

alterava-se substancialmente em relação à primeira avaliação⁽²⁵⁾.

O seguimento (follow-up) final foi superior a 80%?

É inevitável que alguns doentes não terminem o estudo, especialmente se a sua duração for longa. Se o número dos que saem (*drop-out rate*) é muito elevado, então isto constitui um problema, já que se sabe que os doentes que saem têm um prognóstico diferente dos que permanecem no estudo.

Se as razões para o abandono não se relacionarem com o prognóstico, então estas saídas poderão não influenciar os resultados por aí além (para lá da perda de poder e precisão). O contrário verifica-se se elas forem devidas a falta de eficácia do medicamento, a um perfil de efeitos secundários muito desfavorável ou uma adesão particularmente deficiente ao esquema de tratamento, por exemplo⁽¹⁹⁾.

Qual é a dimensão da taxa de abandonos a partir da qual o estudo fica comprometido? Não existe uma resposta simples para esta pergunta, quanto mais pequena melhor. De qualquer maneira, aceita-se que uma taxa superior a 20% é geradora de viéses e indica problemas importantes do estudo.

Importância dos Resultados

A dimensão do efeito terapêutico (RRR, RRA, NNT) foi importante?

A apresentação de resultados de ensaios clínicos deve ser feita de modo *standard*: nos RCTs, os autores habitualmente apresentam proporções (ou taxas, ou percentagens) de eventos (cirrose, hemorragias gastrointestinais, morte) em cada grupo de estudo: experimental ou de controlo. Estes eventos são habitualmente dicotómicos, isto é, caracterizam-se por estarem ou não estarem presentes. A relação entre uma intervenção e o seu resultado é capturada por um conjunto de medidas – designadas como de associação – que incluem o risco absoluto e redução do risco absoluto, o risco relativo e redução do risco relativo, o *odds ratio* e o número necessário tratar. A sua definição consegue-se através da construção das chamadas tabelas de contingência 2 x 2. Um exemplo esquemático é o do Quadro II.

A apresentação dos resultados de uma intervenção num RCT deverá incluir, pelo menos, a RRR, a RRA e o NNT. Se por exemplo só tivermos informação sobre a RRR, então não poderemos conhecer o risco de base (o risco absoluto): uma RRR de 33%, por exemplo, pode representar uma redução do risco de mortalidade de 15% para 10% (uma RRA de 5%, em que só temos de tratar 20 doentes para evitar um evento - NNT=20) ou de 45% para 30% (uma RRA de 15% e em basta tratar 7 - NNT=7).

Quadro II - Uma tabela 2 x 2 e seus resultados derivados.

		Resultado (<i>outcome</i>)	
		SIM	NÃO
Exposição	SIM	A	b
	NÃO	C	d

Risco Relativo. $RR = [a / (a + b)] / [c / (c + d)]$
 Redução do Risco Relativo. $RRR = [c / (c + d)] - [a / (a + b)] / [c / (c + d)]$
 Redução do Risco Absoluto. $RRA = [c / (c + d)] - [a / (a + b)]$
 Número Necessário Tratar. $NNT = 1 / RRA$
 Odds Ratio. $OR = [(a/b) / (c/d)] = ad / cb$

Finalmente, na definição da importância dos resultados, haverá que não esquecer que existe uma distinção entre a significância estatística e a significância clínica. Como já foi discutido, pode existir um resultado estatisticamente significativo que não tem significado clínico.

A estimativa do efeito é suficientemente precisa (IC)?

Na prática, nunca teremos a certeza absoluta de qual é precisamente a redução do risco provocada por uma intervenção. O que podemos esperar são estimativas probabilísticas determinadas pelos ensaios clínicos – as chamadas estimativas pontuais – que, por definição, nunca são exactas ⁽¹²⁾.

Existe um processo de determinar qual é a precisão dos resultados, que é calcular os intervalos de confiança (IC). Como se define um IC? É o intervalo dentro do qual se espera que o valor real de uma medida estatística se situe; é geralmente acompanhado por uma percentagem (quase sempre 95%) que define o nível de confiança respectivo: em 95% das vezes o valor estará dentro dos limites definidos. Quanto menor for o IC, mais precisos são os resultados. Para obtermos um IC com limites estreitos, haverá que aumentar o nº de eventos estudados, que é como quem diz, o nº de doentes estudados. Deste modo, quanto maior for a amostra, mais preciso é o IC.

Esse efeito tem importância clínica?

A resposta a esta questão implica uma avaliação clínica algo subjectiva, que tem a ver com o peso da patologia em questão, o impacto da terapêutica seleccionada, as preferências clínicas, etc. Trata-se acima de tudo de uma abordagem de raciocínio e não estatística ou metodológica.

Aplicabilidade dos Resultados

Os doentes do estudo são semelhantes aos da prática clínica do médico individual?

Um dos aspectos mais importantes dos ensaios clínicos é o da sua aplicabilidade a outros doentes que não os estu-

dados. Muitas vezes o nosso doente tem atributos ou características diferentes dos que participaram no ensaio, pode apresentar diferenças - por exemplo, relativamente à idade, doenças concomitantes, gravidade da doença de base, terapêuticas prévias - que o tornariam inelegível para o estudo, mas que poderão não impedir que os resultados do estudo lhe sejam aplicáveis.

Se o doente fosse elegível para o estudo, ou seja, se o doente preencher todos os critérios de inclusão e não violar nenhum dos critérios de exclusão, pode-se considerar com razoável confiança que os resultados lhe são aplicáveis. Se não é isto que acontece, é necessário estabelecer um julgamento para este doente específico ⁽²⁶⁾.

Foram considerados todos os resultados clínicos importantes?

Em condições ideais, ao tomar decisões sobre tratamentos devem-se procurar ensaios clínicos metodologicamente fortes, que examinem o impacto da terapêutica sobre resultados importantes para o doente, como qualidade de vida ou resultados objectivos (hemorragia gastrointestinal, recidiva de úlcera péptica ou morte, por ex.).

O que é importante para os doentes são os resultados que signifiquem vantagens clínicas: a mortalidade, a qualidade de vida, os efeitos secundários, etc. Mas muitas vezes utilizam-se nos ensaios clínicos os chamados marcadores substitutos (*surrogate end points*), que se definem como uma medida laboratorial ou fisiológica usada como um substituto de um resultado que mede directamente como o doente se sente, funciona ou sobrevive. Os investigadores muitas vezes recorrem a estas variáveis substitutas porque lhes permitem fazer ensaios clínicos em menos tempo e com menor número de doentes (logo, menos dispendiosos). No entanto, só se pode considerar que um marcador substituto é consistentemente seguro quando há uma relação causal entre a alteração desse resultado e a alteração do resultado clinicamente importante. O marcador substituto deve estar integrado na sequência de causalidade do processo patológico, e o efeito de uma intervenção sobre o resultado clínico de interesse deve ser capturado na sua globalidade por uma alteração do marcador substituto.

Os benefícios do tratamento sobrepõem-se aos potenciais riscos e custos da sua implementação?

Um elemento chave na decisão de escolher a terapêutica para um doente é a determinação do risco de um evento adverso nesse doente na ausência de tratamento. As-

sumindo que a RRR é constante ao longo de toda a gama de susceptibilidades, para qualquer dada RRR, quanto mais alta for a probabilidade de que um doente experiente um resultado adverso se não for tratado, maior a probabilidade do doente beneficiar do tratamento e menor o número de doentes análogos que necessitamos de tratar para impedir esse resultado.

O cálculo do NNT específico do doente (ver acima) é uma ajuda importante no processo de pesar os benefícios e os prejuízos associados com as opções de tratamento, mas temos ainda de perceber se a sua implementação é muito complexa (e como tal viável) e qual o seu custo real.

CONCLUSÕES

A análise crítica da evidência científica constitui um passo fundamental na (indispensável) utilização da ciência como base de decisão para a prática clínica. Isto consegue-se através da análise estruturada das publicações, utilizando as respostas a questões metodológicas específicas (as chamadas “guias”) aplicadas a cada estudo, e que são diferentes conforme o tipo de desenho de base da investigação analisada.

Os ensaios clínicos, como instrumentos de determinação do benefício e risco de intervenções terapêuticas (farmacológicas ou outras), constituem peças de evidência absolutamente fundamentais para a prática clínica, necessitando por isso de uma abordagem sistematizada da sua qualidade metodológica e prática.

O esquema apresentado neste artigo poderá ser de grande utilidade para quem queira, de modo individual, analisar e aplicar os dados da investigação na terapêutica.

BIBLIOGRAFIA

1. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. *Clinical Epidemiology*. 2 ed. Boston: Little, Brown and Company; 1991.
2. Wright J, Hill P. *Clinical Governance*. 1st ed. Edinburgh: Churchill Livingstone; 2003.
3. Soares I, Carneiro AV. Princípios de análise metodológica de ensaios terapêuticos. Parte I - Validade dos resultados. *Rev Port Cardiol* 2002; 21: 457-469.
4. Soares I, Carneiro AV. Princípios básicos de análise metodológica de ensaios terapêuticos. Parte II - Importância dos resultados. *Rev Port Cardiol* 2002; 21: 613-626.
5. Soares I, Carneiro AV. Princípios básicos da análise metodológica de ensaios terapêuticos. Parte III - Aplicabilidade dos resultados ao doente individual. *Rev Port Cardiol* 2002; 21: 783-795.
6. Carneiro AV. Princípios básicos de selecção e uso de testes diagnósticos: análise dos resultados dos estudos diagnósticos. *Rev Port Cardiol* 2001; 20: 1135-1139.
7. Carneiro AV. Princípios básicos de selecção e uso de testes diagnósticos: propriedades intrínsecas dos testes. *Rev Port Cardiol* 2001; 20: 1267-1274.
8. Carneiro AV. Princípios de selecção e uso de testes diagnósticos: aplicabilidade prática ao doente individual. *Rev Port Cardiol* 2002; 21: 75-79.
9. Levine M, Walter S, Lee H et al. Users' guides to the medical literature. IV. How to use an article about harm. *JAMA* 1994; 271: 1615-1619.
10. Carneiro AV. Avaliação crítica de artigos sobre prognóstico: regras práticas. *Rev Port Cardiol* 2002; 21: 891-900.
11. Oxman AD, Cook DJ, Guyatt GH, The Evidence-Based Medicine Working Group. Users' guides to the medical literature. VI. How to use an overview. *JAMA* 1994; 272: 1367-1371.
12. User's Guides to the Medical Literature. A Manual for Evidence-Based Clinical Practice. 1st ed. Chicago: JAMA & Archives Journals AMA; 2001.
13. Greenhalgh T. *How to read a paper*. 3d ed. London: BMJ Books; 2006.
14. Altman DG, Schulz KF, Moher D et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Int Med* 2001; 134: 663-694.
15. Carneiro AV. Cálculo da dimensão da amostra em estudos clínicos: princípios metodológicos básicos. *Rev Port Cardiol* 2003; 22: 1513-1521.
16. Matthews JNS. *Introduction to randomized controlled trials*. 2nd ed. Boca Raton: Chapman & Hall/CRC; 2006.
17. Fletcher RH, Fletcher SW. *Clinical epidemiology*. 4th ed. Baltimore: Lippincott Williams & Wilkins; 2005.
18. Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. *Lancet* 2002; 359: 614-618.
19. Simon SD. *Statistical evidence in medical trials*. 1st ed. Oxford: Oxford University Press; 2006.
20. Lang TA, Secic M. *How to report statistics in medicine*. 1st ed. Philadelphia: ACP; 1997.
21. Straus SE, Richardson WS, Glasziou P, Haynes RB. *Evidence-based Medicine. How to practice and teach EBM*. 3d ed. Edinburgh: Elsevier; 2005.
22. Schulz KF, Grimes DA. Sample size calculations in randomised trials: mandatory and mystical. *Lancet* 2005; 365: 24-25.
23. Schulz KF, Grimes DA. Unequal group sizes in randomised trials: guarding against guessing. *Lancet* 2002; 359: 966-970.
24. Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. *Lancet* 2002; 359: 696-700.
25. Gotzsche PC. Blinding during data analysis and writing of manuscripts. *Control Clin Trials* 1996; 17: 285-290.
26. Carneiro AV. Aplicação ao doente individual dos resultados dos ensaios clínicos: regras práticas. *Rev Port Cardiol* 2003; 22: 259-268.