

# Reclamação de Eficácia Terapêutica de um Novo Medicamento na Ausência de Dados Provenientes de Ensaios Clínicos: Problemas e Limitações [85]

ANTÓNIO VAZ CARNEIRO

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

Rev Port Cardiol 2006; 25 (12): 1127-1142

## RESUMO

A definição de eficácia de uma terapêutica cardiovascular baseia-se frequentemente na análise do seu impacto em marcadores substitutivos de doença (MS). Estes definem-se como um resultado laboratorial, um sinal físico ou uma imagem, que são utilizados em ensaios terapêuticos como substituto de marcadores clínicos relevantes e significativos, que representam a maneira como os doentes se sentem, funcionam ou sobrevivem e que, supostamente, predizem o efeito do tratamento.

São por vezes reclamadas eficácias semelhantes em medicamentos dissimilares, com o argumento que ambas têm efeitos semelhantes nos mesmos MS. Esta abordagem envolve uma generalização de uma classe terapêutica para outra ou, dentro da mesma classe terapêutica, de uma molécula para outra, implicando que os benefícios de um medicamento não avaliado serão idênticos aos de outro medicamento cuidadosamente avaliado. Esta generalização é mais usual quando estamos perante 2 medicamentos da mesma classe terapêutica, similares nas suas ações ou nas suas composições químicas, sendo bem mais incerta quando estamos perante dois medicamentos de classes diferentes.

Neste artigo discutimos a legitimidade da reclamação de eficácia e segurança de um (qualquer) novo medicamento cardiovascular, sem o apoio de dados provenientes de ensaios clínicos de grandes dimensões e de longa duração desse medicamento, com o argumento

## ABSTRACT

### **Efficacy claims for new drugs without clinical trial outcome data: problems and pitfalls**

The definition of efficacy of a cardiovascular treatment is frequently based on its impact on surrogate endpoints. These can be laboratory results, physical signs or images that are used in clinical trials as a substitute for clinically relevant endpoints, representing how patients feel, function or survive and that supposedly predict treatment efficacy.

Similar efficacy claims can be made for different drugs, based on their similar effects on the same endpoints. This approach involves generalization from one therapeutic class to another or, within the same class, from one molecule to another, thus implying that the benefits of a non-evaluated drug will be the same as those of another that has been carefully evaluated. Such generalization is more common when dealing with two drugs from the same class that are similar in chemical composition, but is less reliable when the two drugs belong to different classes.

In this article we discuss the legitimacy of efficacy and safety claims for new cardiovascular drugs without outcome data from large clinical trials, based on the similarity of effects found with similar drugs that have clinical trial outcome data.

da similaridade dos efeitos encontrados noutros medicamentos semelhantes (e possuidoras de dados de ensaios clínicos de boa qualidade).

#### Palavras-Chave

Marcadores substitutivos de doença; Ensaio clínico aleatorizado; Prospectivo e controlado; Equivalência terapêutica; Efeito de classe medicamentosa; Cardiologia baseada na evidência

#### Key words

Surrogate endpoints; Randomized clinical trial; Therapeutic equivalence; Drug class; Evidence-based cardiology

## INTRODUÇÃO

A definição da eficácia de uma terapêutica cardiovascular, especialmente no que concerne a modulação de factores de risco (FR), baseia-se frequentemente na análise do seu impacto em marcadores substitutivos de doença (MS - *surrogate endpoints*). Estes definem-se como um resultado laboratorial, um sinal físico ou uma imagem, que são utilizados em ensaios terapêuticos como substitutos de *endpoints* (resultados, marcadores) clínicos relevantes e significativos, que representam a maneira como os doentes se sentem, funcionam ou sobrevivem e que, supostamente, predizem o efeito do tratamento<sup>(1)</sup>.

A utilização permanente e frequente dos MS deve-se ao facto destes diminuírem o custo dos ensaios clínicos (por redução da amostra e do tempo de *follow-up*), assim como evitarem o uso de metodologias demasiado invasivas (e incómodas ou perigosas para o doente) que possam questionar os aspectos éticos do estudo<sup>(2)</sup>.

Muitas vezes são reclamadas eficácias semelhantes em medicamentos dissimilares, com o argumento de que ambas têm efeitos semelhantes em idênticos MS. Esta abordagem envolve generalização de uma classe terapêutica para outra ou, dentro da mesma classe terapêutica, de uma molécula para outra, implicando que os benefícios do medicamento não avaliada serão idênticos aos do medicamento correcta e cuidadosamente avaliado<sup>(3)</sup>. Esta generalização é mais usual quando estamos perante dois medicamentos da mesma classe terapêutica, similares nas suas acções ou nas suas composições químicas, sendo bem mais incerta quando estamos perante medicamentos de classes diferentes.

O problema com o uso de MS como

## INTRODUCTION

The definition of efficacy of a cardiovascular treatment, particularly in terms of modifying risk factors, is frequently based on its impact on surrogate endpoints. These can be laboratory results, physical signs or images that are used in clinical trials as a substitute for clinically relevant endpoints, representing how patients feel, function or survive, and that supposedly predict treatment efficacy<sup>(1)</sup>.

Surrogate markers are commonly used because they can lower the cost of clinical trials by reducing the sample size and follow-up period, as well as avoiding the use of invasive methods that may be unpleasant or dangerous for the patient and thus ethically questionable<sup>(2)</sup>.

Similar efficacy claims are frequently made for different drugs, based on their similar effects on the same endpoints. This approach involves generalization from one therapeutic class to another, or, within the same class, from one molecule to another, thus implying that the benefits of a non-evaluated drug will be the same as those of another that has been properly and thoroughly evaluated<sup>(3)</sup>. Such generalization is more common when dealing with two drugs from the same class that are similar in action and chemical composition, but is less reliable when the drugs belong to different classes.

The problem with the use of surrogate endpoints as markers of therapeutic efficacy is that they are not always valid and reproducible. For example, in the case of hypertension, clinical trials have demonstrated a variable relationship between reductions in blood pressure (the surrogate marker) and clinically relevant events (stroke and coronary artery disease [CAD]), a recent systematic review concluding that although three different

marcadores de eficácia terapêutica é que eles nem sempre são válidos de modo constante e reprodutível: no caso da hipertensão arterial, por exemplo, os ensaios clínicos têm demonstrado uma relação variável entre a redução da tensão arterial (o MS) e os eventos clinicamente relevantes - acidente vascular cerebral (AVC) e doença coronária - tendo uma recente revisão sistemática concluído que, embora em três regimes (doses baixas *versus* altas de diuréticos *versus* beta-bloqueantes) todos baixassem igualmente a tensão arterial, apenas os diuréticos em doses baixas preveniam a DC<sup>(4)</sup>.

Neste artigo iremos discutir em detalhe a legitimidade da reclamação de eficácia e segurança de um (qualquer) novo medicamento cardiovascular, sem o apoio de dados provenientes de ensaios clínicos de grandes dimensões e de longa duração desse medicamento, com o argumento da similaridade dos efeitos encontrados noutros medicamentos semelhantes (e possuidoras de dados de ensaios clínicos de grande qualidade). Começaremos por analisar o ensaio clínico aleatorizado, prospectivo e controlado (*randomized clinical trials - RCTs*) como a metodologia científica mais eficaz para a prova de eficácia de uma qualquer intervenção (incluindo farmacológica); de seguida, afloraremos a maneira de aplicar os resultados de RCTs a populações diversas da amostra estudada; depois, discutiremos as vantagens e desvantagens dos MS de doença. Apresentaremos então a definição rigorosa de equivalência terapêutica e o conceito de efeito de classe medicamentosa, terminando com uma síntese sobre os problemas da reclamação de eficácia medicamentosa, na ausência de RCTs de boa qualidade.

### **O ENSAIO CLÍNICO COMO METODOLOGIA CIENTÍFICA DE EFICÁCIA TERAPÊUTICA E PREVENTIVA**

Quando um médico administra uma qualquer terapêutica medicamentosa a um doente, baseia-se em quatro princípios essenciais: 1) o diagnóstico da doença está bem estabelecido; 2) o medicamento selecionado é eficaz e seguro para a patologia em causa, baseado em dados

treatment regimens (low- vs. high-dose diuretics vs. beta-blockers) all lowered blood pressure to the same extent, only low-dose diuretics reduced the risk of CAD<sup>(4)</sup>.

In this article we discuss in detail the legitimacy of efficacy and safety claims for new cardiovascular drugs without outcome data from large clinical trials with a long follow-up, based on the similarity of effects found with similar drugs that have clinical trial outcome data. We will begin by analyzing prospective randomized controlled trials (RCTs), which are the most effective method of proving the efficacy of any intervention, including pharmacological therapy, and touch on how to apply the results of RCTs to populations that differ from the trial sample. The advantages and disadvantages of surrogate disease markers will then be discussed, as well as the strict definition of therapeutic equivalence and the concept of drug class effect. Finally, we will summarize the problems of efficacy claims without outcome data from high-quality RCTs.

### **THE CLINICAL TRIAL AS A METHOD OF ASSESSING THERAPEUTIC AND PREVENTIVE EFFICACY**

When physicians prescribe a drug for their patients, their decision is based on four principles: 1) the diagnosis of the disease is clearly established; 2) the drug is effective and safe for the pathology in question, preferably based on data from RCTs; 3) the choice of a specific drug is made bearing in mind all the available alternative drugs in the same class; and 4) application of the therapy is based on a set of guidelines that enable the results of RCTs to be applied to the individual patient. Points 3) and 4) will be discussed later; point 2) is discussed here.

The choice of randomized controlled trials as the best evidence on therapeutic (or preventive) measures is based on the fact that this design is the only one that allows the establishment of causality, i.e. a relationship between an intervention and an outcome<sup>(5-7)</sup>. Testing a hypothesis is the first step in the scientific method, since only experiment can, within certain limits, ensure detection of a real effect, whose size (presented as relative risk reduction, absolute risk reduction or number needed to

provenientes (preferencialmente) de RCTs; 3) a selecção do medicamento específico foi feita tendo em conta as alternativas disponíveis na mesma classe medicamentosa; e 4) a aplicação desta terapêutica ao doente assenta num conjunto de regras próprias, codificadas, que permitem transferir os resultados dos RCTs ao doente individual. Os pontos 3) e 4) serão discutidos mais adiante, iremos discutir aqui o ponto 2).

A razão para a escolha de ensaios clínicos aleatorizados, prospectivos e controlados como evidência de alta qualidade sobre terapêutica (ou prevenção), provém do facto deste desenho de estudo ser o único que permite um estabelecimento rigoroso da causalidade, isto é, da relação entre uma intervenção e um resultado<sup>(5-7)</sup>. Isto tem a ver com o facto de que testar uma hipótese constitui o primeiro passo do processo científico, já que a metodologia experimental é a única que pode garantir - dentro de certos limites - a detecção de um efeito verdadeiro, cuja dimensão (apresentada como redução do risco relativo, redução do risco absoluto ou número necessário tratar) e precisão (calculada através dos valores *P* ou dos intervalos de confiança) podem ser calculadas de modo preciso. Para isso, os investigadores utilizam métodos estatísticos para testar a hipótese nula (que não há diferenças entre dois tratamentos), observando se os dados obtidos no ensaio clínico são ou não consistentes com aquela.

Para a definição de um benefício terapêutico, necessitamos de marcadores de doença, isto é, características ou factores que mantêm uma relação estreita com a patologia em causa e que podem ser mensuráveis de modo rigoroso e reproduzível. O quadro I descreve alguns marcadores de eficácia passíveis de ser utilizados em RCTs<sup>(8)</sup>.

#### QUADRO I

##### Alguns marcadores para detecção de eficácia terapêutica em RCTs

- Tempo (dias, meses) e percentagem (75%, 100%, etc.) de melhoria de uma determinada variável
- Intervalo até recorrência de sintomas após paragem do tratamento (retorno a 75% ou 80% da linha de base, por ex.)
- Intervalo de tempo, sob tratamento activo, até um novo episódio da doença aparecer
- Grau de recorrência de sintomas após cessação do tratamento

treat) and precision (determined by p values or confidence intervals) can be calculated accurately. For this purpose, investigators use statistical methods to test the null hypothesis (that there is no difference between two treatments) by analyzing whether it is consistent with the data obtained in the trial.

In order to determine therapeutic benefit, endpoints are required, in other words, characteristics or factors that are closely linked to the pathology in question and can be measured in a precise and reproducible manner. *Table I* describes some markers of efficacy that may be used in RCTs<sup>(8)</sup>.

TABLE I

##### Some markers of therapeutic efficacy in RCTs

- Time (days, months) and percentage (75%, 100%, etc.) of improvement in a given variable
- Time interval of recurrence of symptoms after end of treatment (e.g. return to 75% or 80% of baseline)
- Time interval under active treatment before a new episode of the disease
- Degree of recurrence of symptoms after end of treatment
- Duration of clinical improvement under maintenance treatment
- Magnitude of the treatment effect at specified time intervals after the beginning of treatment (e.g. 7 days, 30 days, 180 days)
- Indices of efficacy calculated on the basis of surrogate markers (ejection fraction, end-diastolic ventricular pressure, etc.)
- Subjective indices of the patient's improvement
- Indices related to quality of life
- Improvement or disappearance of risk factors for the disease
- General and specific mortality of the disease
- Morbidity (adverse reactions, mean time of hospitalization, etc.)
- Overall clinical impression of the patient's state of health as assessed by doctor or patient
- Time interval between episodes of the disease (e.g. convulsions)
- Study rate dropout
- Need for concomitant therapy (e.g. anti-inflammatory drugs in osteoarthritis)
- Prevention of physical deterioration from the disease

To determine the efficacy of a drug through an RCT, a clearly defined methodology is required in order to ensure that the effects observed are real and clinically important and are generalizable.

- Duração da melhoria clínica sob tratamento de manutenção
- Magnitude do efeito terapêutico em intervalos de tempo determinados após início do tratamento (por ex. 7 dias, 30 dias, 180 dias)
- Parâmetros de eficácia calculados com base em marcadores substitutivos (fracção de ejeção, pressão ventricular no fim da diástole, etc.)
- Parâmetros subjectivos de melhoria do doente
- Parâmetros relacionados com qualidade de vida
- Melhoria ou desaparecimento de factores de risco da doença
- Mortalidade geral e específica de doença
- Morbilidade (reações adversas, demora hospitalar média, etc.)
- Impressão clínica global do estado de saúde do doente analisado pelo médico ou pelo doente
- Intervalo de tempo entre episódios da doença (convulsões, por ex.)
- Taxa de *drop-outs* do estudo
- Necessidade de terapêutica concomitante (anti-inflamatórios na osteoartrite, por ex.)
- Prevenção da deterioração física provocada pela doença

No caso da determinação da eficácia de um medicamento através de um RCT, é necessária uma metodologia muito precisa, que garanta que os efeitos observados são reais, clinicamente importantes e generalizáveis. Isto consegue-se agrupando indivíduos seleccionados de uma população em estudo (geralmente porque sofrem de uma patologia de interesse), que são distribuídos aleatoriamente em dois (por vezes mais) grupos ou braços: experimental (expostos a uma intervenção) e controlo (submetidos por norma a um placebo ou a uma medicação convencional)<sup>(5)</sup>. A razão para utilizar esta metodologia é permitir uma distribuição equitativa de factores que podem influir no prognóstico, equilíbrio que é considerado essencial na apreciação de uma relação de causalidade entre uma intervenção e o efeito observado. Os resultados deste tipo de estudo são avaliados comparando os grupos em termos de taxas de eventos (*ver quadro acima*)<sup>(9)</sup> e, idealmente, serão aplicáveis a todos os outros doentes similares.

This is achieved through randomly dividing individuals selected from the study population (generally because they have the disease under study) into two or more groups or arms: an experimental group (exposed to the intervention) and a control group (usually given a placebo or the standard drug)<sup>(6)</sup>. The reasoning behind this approach is to ensure an equal distribution of factors that may influence prognosis, which is essential to determine a cause-effect relationship between an intervention and the effect observed.

Results from this type of study are assessed by comparing event rates in the different groups (see above table)<sup>(9)</sup> and, ideally, will be applicable to all similar patients.

### **APPLICATION OF CLINICAL TRIAL OUTCOME DATA TO POPULATIONS THAT DIFFER FROM THOSE STUDIED**

One of the most important aspects of clinical trials is the applicability of the results to populations and/or individual patients that differ from those assessed in a particular trial; this is termed the external validity or generalization of results of an RCT<sup>(10)</sup>.

The question of external validity is extremely important for any physician when treating patients, but is even more so for a pharmaceutical company that wants to market a drug or gain approval for its use in national health systems, since it has to prove that the results obtained in an RCT population are applicable to patients in the real world who may be prescribed the drug once it is available. Evidence that the benefits demonstrated in a clinical trial can be generalized is absolutely essential in this context.

Factors that can affect the external validity of an RCT include<sup>(11)</sup>: the context in which the trial was carried out, sample selection and patients' clinical characteristics, differences between the study protocol and current clinical practice, and patient outcomes and follow-up, as well as different rates of side effects. Results can also be strongly influenced by the doctor-patient relationship, the placebo effect, and patient preferences.

To ensure the applicability of clinical trial results, an affirmative response should be obtained to the following questions<sup>(12)</sup>:

- Were the patients in the trial similar to our

## A APLICAÇÃO DOS RESULTADOS DOS ENSAIOS CLÍNICOS A POPULAÇÕES DIFERENTES DA ESTUDADA

Um dos aspectos mais importantes no que concerne aos ensaios clínicos é a aplicabilidade dos seus resultados a populações e/ou doentes individuais, diferentes dos analisados nesse ensaio específico - a chamada validade externa ou generalização dos resultados do RCT<sup>(10)</sup>.

Esta questão da validade externa é muito importante para o médico responsável pelo tratamento de doentes, mas é-o ainda mais para a indústria farmacêutica, quando quer ver o seu medicamento introduzido no mercado ou participado, tendo de provar que os resultados obtidos numa amostra de um RCT são aplicáveis a outros doentes (os doentes "reais" que irão ser sujeitos à medicação, uma vez esta disponível para prescrição). A prova da existência de generalização dos benefícios encontrados num ensaio clínico é, neste contexto, absolutamente crucial.

Os factores que podem afectar a validade externa de um RCT incluem (entre outros)<sup>(11)</sup>: o contexto em que se realizou o ensaio, a selecção amostral dos doentes e suas características clínicas, as diferenças entre o protocolo do estudo e a prática clínica corrente, os resultados (*outcomes*) e o seguimento (*followup*) dos doentes, assim como taxas diversas de efeitos secundários. Os resultados podem ainda ser poderosamente influenciados pela relação médico-doente, pelo efeito placebo e pelas preferências dos doentes.

Para garantir a aplicabilidade dos resultados do ensaio clínico, deveremos obter respostas afirmativas às seguintes perguntas<sup>(12)</sup>:

- Os doentes do estudo eram semelhantes ao doente que queremos tratar?
- Foram considerados todos os resultados clinicamente relevantes?
- Os benefícios que o tratamento pode trazer ao doente valem os efeitos adversos potenciais e os custos?
- Quais são os valores e expectativas do nosso doente em relação aos resultados positivos e adversos do tratamento que lhe estamos a oferecer?

Como já discutimos em artigo anterior publicado na Revista Portuguesa de Cardiologia<sup>(7)</sup>, para resolver este problema a solução mais simples seria encontrada na resposta à questão:

patient?

- Were all the results considered clinically relevant?
- Do the benefits that the treatment may bring to the patient outweigh its possible adverse effects and the cost?
- What are our patient's values and expectations with regard to the positive and adverse results of the treatment that is being offered?

As already discussed in a previous article published in the *Portuguese Journal of Cardiology*<sup>(7)</sup>, the simplest solution to this problem is to be found in the answer to the question: "Could the patient have been enrolled in the study by satisfying its inclusion and exclusion criteria?". An affirmative response implies accepting the applicability of the trial's findings; however, a negative response (because the patient presents very different demographic features or serious comorbidities not found in the trial subjects, for example) would compromise to a varying extent the transfer of the RCT's therapeutic results to the individual patient.

Clinical trials cannot produce results that would be relevant for all potential candidates for a particular treatment, but to be considered valid, they must be presented in such a way that clinicians can easily judge which patients are eligible and which clinical markers should be assessed.

## SURROGATE ENDPOINTS AS EVIDENCE OF THERAPEUTIC EFFICACY

Among the various methodological components of an RCT, the endpoints, also known as disease markers, are of particular importance. The whole purpose of an RCT is to measure changes in the disease process following a therapeutic or preventive intervention, and specific measures of such changes must be predefined - the endpoints.

As stated above, problems arise when assessment of therapeutic efficacy is to be based on surrogate endpoints rather than actual disease markers. There are various disadvantages to using surrogate endpoints in therapeutic trials and appraising them correctly can be complicated: for example, a



“Será que o doente poderia ter participado - satisfazendo os critérios de inclusão e exclusão - no referido estudo?”. Uma resposta afirmativa implicaria a aceitação da aplicabilidade dos achados do estudo, mas uma resposta negativa (porque o doente apresenta características demográficas muito diversas ou patologias concomitantes graves não presentes nos sujeitos do estudo, por exemplo), comprometeria em grau variável a transferência dos dados terapêuticos do RCT para o doente individual.

Os ensaios clínicos não podem produzir resultados que sejam relevantes para a totalidade dos doentes potencialmente candidatos para tratamento mas, para poderem ser considerados válidos, devem ser apresentados de maneira a que os clínicos possam julgar com clareza quais os doentes elegíveis e que marcadores clínicos se devem ter em conta.

### MARCADORES SUBSTITUTIVOS COMO PROVA DE EFICÁCIA TERAPÊUTICA

De entre os diversos componentes metodológicos de um RCT, os marcadores de doença (designados indiferentemente por *endpoints* ou *disease markers*) revestem-se de particular importância. Com efeito, e em última análise, o que se pretende num RCT é precisamente medir alterações no processo de doença pela intervenção terapêutica (ou preventiva), sendo que essas alterações têm de ser definidas previamente através de medidas específicas - os tais marcadores de doença.

Como já foi dito, o problema surge quando se pretende utilizar - como base de avaliação de eficácia terapêutica - não os marcadores de doença, mas sim os seus substitutos. Os obstáculos na utilização de MS em ensaios terapêuticos são de várias ordens e a sua correcta avaliação apresenta alguma complexidade: por ex. uma mudança num MS pode não responder às questões sobre o objectivo e a selecção da melhor terapêutica para uma determinada condição; o MS pode também não ser válido ou fiável na patologia em estudo e o uso de um MS isolado ignora outras medidas potencialmente importantes de eficácia terapêutica.

Para que um MS possa ser um factor válido na análise de eficácia terapêutica - deve possuir idealmente as características incluídas no quadro II<sup>(6)</sup>.

change in a surrogate endpoint may not answer questions about the therapeutic goal and how to select the best treatment for a particular condition; the surrogate endpoint may not be valid or reliable for the pathology being studied; and use of a single surrogate endpoint ignores other potentially important measures of therapeutic efficacy.

In order for a surrogate endpoint to be a valid factor when analyzing therapeutic efficacy it should ideally possess the characteristics listed in *Table II*<sup>(6)</sup>.

TABLE II

#### Some markers of therapeutic efficacy in RCTs

- It should be reliable, reproducible, clinically available, easily quantifiable, affordable, and show a dose-response effect (the higher the level of the surrogate endpoint, the greater the probability of disease)
- It should be a true predictor of disease (or risk of disease) and not express exposure to any other covariable; the relation between the surrogate endpoint and the disease should make sense in biological terms
- It should have high sensitivity (a positive result for the surrogate endpoint should identify most patients with increased risk of the disease)
- It should have high specificity (a negative result should exclude most patients who do not present increased risk of the disease)
- It should show a clear distinction between normal and abnormal values
- It should have an acceptable positive predictive value (positive results should consistently identify patients with increased risk of the pathology in question)
- It should have an acceptable negative predictive value (negative results should consistently identify patients with reduced risk of the pathology in question)
- It should be amenable to quality control
- Changes in surrogate endpoint levels should rapidly and accurately reflect the response to treatment - for example, levels should normalize in states of remission or cure

In most cases, a surrogate endpoint is a risk factor for the disease, in other words it is one of the causes of the pathology. Its definition is based on two steps: firstly, high-quality observational studies have determined that it is a risk factor for the disease (establishing an association between them), and subsequently RCTs have confirmed that a decrease in the risk factor leads to a parallel reduction in the incidence of the

## QUADRO II

### Características ideais de um marcador substitutivo

- deve ser fiável, reprodutível, clinicamente disponível, facilmente quantificável, acessível e exibindo um efeito de “dose-resposta” (quanto maior for o nível do MS, maior a probabilidade de doença)
- deve ser um preditor verdadeiro de doença (ou risco de doença) e não exprimir uma exposição a uma qualquer outra co-variável; a relação entre o MS e a doença deve fazer sentido sob o ponto de vista biológico
- deve possuir sensibilidade elevada (um resultado positivo do MS deve identificar a maior parte dos doentes com risco aumentado de doença)
- deve possuir especificidade elevada (um resultado negativo deve excluir a maior parte dos pacientes que não apresentam aumento do risco de doença)
- deve exibir um claro nível de definição entre valores normais e anormais
- deve possuir um aceitável valor preditivo positivo (resultados “positivos” devem consistentemente identificar doentes com risco aumentado para a patologia em causa)
- deve possuir um aceitável valor preditivo negativo (resultados “negativos” devem consistentemente identificar doentes com risco diminuído para a patologia em causa)
- deve ser passível de controlo de qualidade
- mudanças nos níveis do MS devem reflectir (de maneira rápida e precisa) a resposta ao tratamento - por exemplo, os níveis devem normalizar em estádios de remissão ou cura

Na maior parte dos casos, um MS é importante acima de tudo como um factor de risco de doença, isto é, como um factor causal dessa patologia. Esta definição constrói-se em dois passos: inicialmente, estudos observacionais de boa qualidade definem o MS como um FR de doença (associação entre eles) e, posteriormente, RCTs confirmam a diminuição paralela deste FR e da incidência da doença (relação causal entre ambos)<sup>(13)</sup>. Mas atenção: os RCTs não são grande coisa para testar FR ou mecanismos de doença, porque o que eles numa maneira geral testam são medicamentos específicos em dosagens determinadas. Assim, um ensaio clínico com um medicamento de uma determinada classe não fornece grande informação sobre a validade do MS, já que a eficácia da medicamento pode ser devida a um mecanismo totalmente independente daquele (daí a necessidade da realização de vários RCTs com várias medicamentos medindo o mesmo MS)<sup>(3)</sup>.

Mas é claro que a utilização mais importante

disease (thus establishing a causal relationship between the two)<sup>(13)</sup>. However, caution is required, since RCTs are not generally designed to test disease mechanisms or risk factors, but rather specific drugs at particular dosages. Thus, a clinical trial testing a drug of a particular class does not provide much information on the validity of the surrogate endpoint, since the drug's therapeutic efficacy may be due to a totally independent mechanism. Various RCTs with different drugs measuring the same surrogate endpoint are therefore required<sup>(3)</sup>.

Nevertheless, the most important use of surrogate endpoints is clearly in assessing the efficacy and safety of a new treatment, compared to one that is already established. The theoretical basis for this approach is as follows: 1) in a large-scale RCT over a long period, it is found that drug X leads to a reduction in both the surrogate endpoint and the incidence of the disease under study, thus validating the endpoint as a specific marker of the disease; 2) in trials of short duration, drug Y is also found to reduce this endpoint; and so 3) in clinical practice, it can be concluded that drug Y will behave like drug X in terms of the disease's incidence, and is thus a scientifically proven alternative in this setting and can be used in the same way.

The problem with this approach is that there is no empirical evidence to support it. There may be various combinations of risk factors, the new drug may have different side effects or an as yet unknown action, and the duration of trials can vary. Thus, to claim therapeutic equivalence based on the above reasoning would seem unacceptable. As mentioned earlier, this has been observed in clinical trials on antihypertensive therapies, in which lowering blood pressure has not always shown a fixed relationship with clinically relevant events<sup>(4)</sup>.

To overcome these problems, equivalence trials are needed.

## EQUIVALENCE TRIALS

RCTs are crucial for the purposes of licensing new drugs, since only those that have been shown to improve patients' symptoms and/or prognosis will be approved for general use.

Clinical trials comparing a new drug with an



dos MS é na avaliação de eficácia e segurança de uma nova terapêutica, quando comparada com uma já estabelecida. A abordagem teórica envolve o seguinte raciocínio: 1) num RCT de grandes dimensões e de longa duração, verifica-se com a utilização do medicamento X uma redução simultânea do MS e da incidência da doença respectiva (efectivamente validando este como marcador específico da doença); 2) em ensaios de curta duração, o medicamento Y diminui também efectivamente o MS em causa; logo, 3) na prática clínica, pode-se concluir que o medicamento Y ir-se-á comportar como o X no que concerne à incidência de doença, sendo uma alternativa cientificamente provada neste contexto e podendo utilizar-se clinicamente de modo análogo. O problema com esta abordagem é que não existe evidência empírica que a suporte, já que podem existir diversas combinações de FR, os efeitos adversos podem ser muito diversos, o novo medicamento pode ter uma acção ainda desconhecida e a duração dos ensaios ser diversa. Deste modo, reclamar equivalência terapêutica baseada na analogia acima descrita parece ilegítimo. Isto mesmo se verifica nos ensaios clínicos com medicamentos anti-hipertensivos, por exemplo, em que, como já foi dito, a baixa da tensão arterial nem sempre tem demonstrado uma relação fixa com os eventos clinicamente relevantes<sup>(4)</sup>.

Para ultrapassar este dilema necessitamos de ensaios de equivalência entre fármacos.

## ENSAIOS DE EQUIVALÊNCIA DE FÁRMACOS

Hoje em dia os RCTs são cruciais para efeitos de licenciamento de novos medicamentos, já que só serão aprovados para uso generalizado os que tenham provas de melhorar os sintomas dos doentes, o respectivo prognóstico ou ambos.

Os RCTs em que se compara um medicamento (novo) contra um comparador activo são designados “de superioridade”, porque se vai tentar provar que o novo medicamento é, em termos de eficácia, superior ao clássico, isto é, os investigadores estão acima de tudo interessados em descobrir se um tratamento novo reduz a frequência de um dado outcome (resultado, marcador de doença, etc.) prejudicial, quando comparado com outro tratamento já utilizado.

active comparator are called “superiority trials”, since their aim is to prove that the new drug’s efficacy is superior to the standard therapy. The investigators are interested above all in determining whether the new treatment reduces the frequency of a given adverse outcome or disease marker compared to an established treatment.

Efficacy claims between two drugs, only one of which has high-quality RCT data, can be legitimized through equivalence trials. These are designed to prove pharmacological and clinical equivalence between two different active substances, or two different formulations of the same substance. Since it is impossible to prove that two interventions are exactly equivalent, these trials are designed in such a way as to be able to prove that the treatments do not differ by more than a predefined “margin of equivalence”. This type of trial can be valuable when the standard therapy has been shown to be efficacious, but the new treatment is safer, easier to administer, or less expensive<sup>(14)</sup>. In addition, in cases of serious disease for which there is an effective treatment, it would be unethical to compare it with a placebo or a substance known to be less effective.

*Table III* describes the methodological guidelines that should be followed for an equivalence trial to be considered valid<sup>(15)</sup>.

TABLE III

### Guidelines for the critical appraisal of equivalence trials

1. Was the allocation of patients to the treatments randomized?
2. Was the randomization concealed?
3. What was the dropout rate at the end of the trial?
4. Was the analysis performed on the basis of “intention-to-treat” or “on-treatment”?
5. Was treatment blinded to both patients and clinicians?
6. Did both groups receive the same interventions (apart from the experimental treatment)?
7. Were the groups similar at the beginning of the study in terms of known prognostic factors?
8. Were the trial’s endpoints clinically important?
9. Was the active control previously shown to be effective?
10. Were patients and outcomes similar to those of the original studies that demonstrated the efficacy of the active control?
11. Were the two therapeutic regimens applied properly?

A reclamação de eficácia entre dois medicamentos dos quais apenas um possui RCTs de boa qualidade seria legítima através de ensaios de equivalência. Estes destinam-se a provar a equivalência farmacológica e clínica entre duas substâncias activas diferentes, ou da mesma substância com duas formulações galénicas diferentes, por exemplo. Porque é impossível provar que duas intervenções são exactamente equivalentes, os ensaios de equivalência são desenhados de modo a que se possa provar que aquelas não diferem mais do que uma certa “margem de equivalência” previamente definida. Este tipo de ensaios pode ser valioso no caso da terapêutica clássica ter provas de eficácia, mas o novo tratamento ser mais seguro, mais fácil de usar, ou menos dispendioso<sup>(14)</sup>. Por outro lado, nos casos de doenças graves com tratamento eficaz, não é ético comparar um placebo ou uma substância que se sabe ser menos eficaz com esse tratamento.

O quadro III descreve as regras a que deve obedecer, metodologia para que um ensaio de equivalência seja considerado válido<sup>(15)</sup>.

#### QUADRO III

##### Regras para avaliação crítica de ensaios de equivalência

1. a distribuição dos doentes pelos tratamentos foi aleatorizada?
2. a aleatorização foi ocultada?
3. qual foi a taxa de *drop-outs* no final do estudo?
4. a análise foi feita com “intenção-de-tratar” e “em tratamento”?
5. houve ocultação do tratamento para os doentes e os clínicos?
6. os dois grupos sofreram as mesmas intervenções (para além das terapêuticas experimentais)?
7. os grupos foram semelhantes no início, em termos de factores prognósticos conhecidos?
8. os *endpoints* do estudo foram clinicamente importantes e relevantes?
9. existia previamente prova de eficácia do controle activo?
10. os doentes e os *outcomes* foram semelhantes aos dos estudos originais que provaram eficácia do controle activo?
11. os dois esquemas terapêuticos foram aplicados correctamente?
12. a análise estatística foi a correcta para a questão?

12. Was the statistical analysis appropriate for the question posed?
13. Was the equivalence margin specified before the beginning of the trial?
14. Was the trial size sufficient to test the initial hypothesis?

In the absence of equivalence trials, which are still uncommon, drugs can only be compared through indirect evidence of a class effect, but this is considerably less reliable. This means that caution is required in extrapolating RCT data on one drug to another similar drug.

The concept of a drug class effect, which has many advantages but also some problems, will be discussed below.

### THE CONCEPT OF CLASS EFFECT

Drugs that are similar in their action are often used interchangeably, for example beta-blockers in hypertension, ACE inhibitors in heart failure, or statins in dyslipidemia.

There are three elements in the definition of a drug class<sup>(16)</sup>: a similar chemical structure (such as the dihydropyridine ring in some calcium channel blockers); a similar mechanism of action (for example the adrenergic receptor blocking action of beta-blockers); and a comparable pharmacological effect (such as antihypertensive or antianginal drugs).

The assumption that drugs within the same class will exert similar effects (unless, of course, there is clear evidence of important differences) can lead to two errors in extrapolation with possible clinical consequences (for a more detailed discussion, the reader can consult one of our previous articles published in the *Journal*)<sup>(17)</sup>. Firstly, when a class of drugs (such as the thiazide diuretics) produce similar pharmacological effects (blood pressure lowering) and similar clinical effects (stroke reduction), a second class of drugs (for example, the calcium channel blockers) that produce the same pharmacological effect might be assumed to produce the same clinical benefit. In the absence of RCTs verifying that final step, this type of extrapolation may be in error. Even

13. a margem de equivalência foi especificada antes do início do estudo?
14. a dimensão do estudo foi suficiente para testar a hipótese inicial?

Na ausência de ensaios de equivalência (que de resto são ainda muito raros), a comparação entre dois medicamentos só pode ser definida através de provas indirectas, que definam efeitos de classe (mas de modo substancialmente menos seguro). Isto é, a extrapolação de dados de um RCT com um determinado fármaco deve ser feita com cuidado para outro fármaco semelhante.

Entra aqui o conceito de classe terapêutica, que traz consigo muitas vantagens mas também alguns problemas.

### O CONCEITO DE EFEITO DE CLASSE MEDICAMENTOSA

A utilização de medicamentos julgadas similares na sua acção - e portanto substituíveis - é muito frequente: são exemplos desta realidade terapêutica a utilização de beta-bloqueantes como anti-hipertensivos, ou de inibidores da enzima de conversão (IECA) na insuficiência cardíaca congestiva, ou ainda a utilização de estatinas na dislipidémia.

A definição de classe medicamentosa passa pela existência de três conceitos fundamentais<sup>(16)</sup>: existência de uma estrutura química semelhante (por exemplo, o anel dihidropiridina em certos bloqueadores dos canais de cálcio), de um mecanismo de acção semelhante (por exemplo, os beta-bloqueantes possuem uma actividade bloqueadora dos receptores adrenérgicos) ou de um efeito farmacológico sobreponível (por exemplo, anti-hipertensores ou anti-anginosos).

A assunção de que fármacos dentro da mesma classe exercem efeitos semelhantes (a não ser que haja evidência clara da existência de diferenças importantes), pode conduzir a dois erros de extrapolação, com potenciais consequências clínicas - o leitor mais interessado poderá consultar artigo nosso publicado na RPC<sup>(17)</sup>: em primeiro lugar, quando uma classe de fármacos (como os diuréticos tiazídicos) produzem efeitos farmacológicos semelhantes (baixa da tensão arterial) e efeitos clínicos semelhantes (redução do número de

within the same class, individual drugs may have different physiologic effects other than the mechanism of action which defined them as being from the same class. For example, some calcium channel blockers have unfavorable effects on total mortality<sup>(18)</sup>, and while all statins act on the HMG-CoA reductase enzyme, they may have different non-lipid effects on the atherothrombotic process which may influence their clinical efficacy.

*Table IV* shows the hierarchy of levels of evidence to be applied when comparing one drug with others of the same class<sup>(17, 19)</sup>, and lists the potential problems of ascribing a class effect when the available evidence is of variable quality and robustness.

It is also a well-known fact that drugs within the same class can have different safety profiles, which is a crucial consideration when assessing a class effect. A similar approach to that required for efficacy is vital when it comes to safety, and so it is essential to determine the quality of the evidence on which safety claims are based (*Table V*)<sup>(17, 19)</sup>.

As in all approaches that use levels of evidence as the basis for evaluating the intrinsic validity of conclusions, when assessing the efficacy (*Table IV*) and safety (*Table V*) of drugs within the same class, it is accepted that in the absence of the best evidence, one can and should consider lower levels, while bearing in mind the methodological problems that this may involve.

### CONCLUSIONS: ARE EFFICACY CLAIMS LEGITIMATE WITHOUT CLINICAL TRIAL OUTCOME DATA?

In the light of the above, it would seem prudent to base approval of new drugs primarily on validated surrogate endpoints, in order to avoid the error of claiming efficacy and safety through analogy with other drugs that have been thoroughly studied in large-scale clinical trials.

An intervention directed at a surrogate endpoint only produces direct benefit for the patient if it results in or predicts a favorable outcome, such as myocardial infarction (MI) or heart failure (HF). Surrogate markers may not be valid and an unexpected adverse effect of

QUADRO IV

Níveis de evidência para comparação da eficácia dentro de uma mesma classe

| Nível | Comparação  | Doentes do estudo  | Resultados   | Ameaças à validade   |
|-------|---|--|--|--|
| 1     | Dentro de uma RCT fazendo comparação directa entre fármacos   | Recrutados da mesma população (por definição)                          | Resultados importantes para o doente*                                    | Aleatorização não ocultada<br>Ausência de ocultação<br>Follow-up incompleto  |
| 2     | Dentro de um RCT em paralelo  | Idênticos (por definição)  | Resultados substitutos validados†  | As de nível 1<br>Validade do resultado substituto para resultados clinicamente importantes   |
| 2     | Entre RCTs de diferentes fármacos versus placebo  | Semelhantes ou diferentes (em estado de doença e de factores de risco) | Resultados importantes para o doente ou resultados substitutos validados | As de nível 1<br>Diferenças entre os ensaios em:<br>• Características dos doentes<br>• Qualidade metodológica (“blinding” adequado, aleatorização ocultada, etc)<br>• Definições de resultados<br>• Taxas de aderência                             |
| 3     | Entre análises de subgrupo de RCTs de diferentes fármacos vs placebo                                      | Semelhantes ou diferentes  | Resultados importantes para os doentes ou resultados substitutos         | As de nível 1 e 2<br>Múltiplas comparações, após colheita de dados <i>ad hoc</i><br>Subgrupos com poder limitado<br>Classificação incorrecta em subgrupos  |
| 3     | Entre RCTs de diferentes fármacos vs placebo  | Semelhantes ou diferentes  | Resultados substitutos não validados                                     | Os resultados substitutos podem não capturar todos os efeitos (benéficos ou prejudiciais) de um agente terapêutico   |
| 4     | Entre estudos não aleatorizados (estudos observacionais e investigação de bases de dados administrativas) | Semelhantes ou diferentes  | Resultados importantes para os doentes                                   | Confusão/erro por indicação, aderência, ou calendário<br>Factores de confusão desconhecidos/não avaliados<br>Erro de medição<br>Para investigação de resultados: bases de dados limitadas, sistemas de codificação não adequados para investigação |

\*Resultados importantes para o doente referem-se a dados de eficácia a longo prazo e os marcadores específicos dependem da condição em tratamento. No caso das estatinas, que são usadas para prevenir e tratar doença aterosclerótica, os resultados importantes para os doentes incluem mortalidade de todas as causas, enfarte do miocárdio e acidente vascular cerebral.

† Resultados substitutos são considerados validados apenas quando a relação entre o resultado substituto e os resultados importantes para o doente tiver sido firmemente estabelecida em RCTs a longo prazo.

Quadro adaptado de <sup>(19)</sup>.

AVC) pode-se ser levado a considerar que uma segunda classe de fármacos (por exemplo, bloqueadores dos canais de cálcio), que produzem os mesmos efeitos farmacológicos (baixa da tensão arterial), também deverá produzir o mesmo benefício clínico. Este tipo de extrapolação pode ser problemática na ausência de ensaios aleatorizados e controlados que verifiquem este passo final, porque, mesmo no interior de uma mesma classe, diferentes fármacos podem ter diferentes efeitos fisiológicos, para além do mecanismo de acção que serviu para os classificar como pertencendo a essa

a drug may lead to an unfavorable outcome. Such endpoints are involved in the causal chain of events to varying degrees: cholesterol (a biochemical variable), blood pressure (a pathophysiological variable) and coronary artery diameter (a morphological variable) are all surrogate markers for MI and HF, although the third one is the most important as it is most closely linked to cardiovascular events <sup>(20)</sup>. The strength of conviction concerning the validity of a surrogate endpoint to assess therapeutic benefit is to a certain extent related to the degree of biological plausibility

TABLE IV

Levels of evidence for comparing the efficacy of drugs within the same class

| Level | Comparison  | Study patients   | Outcomes  | Threats to validity   |
|-------|---|--|---|---|
| 1     | Within a “head-to-head” RCT   | Identical (by definition)                                | Clinically important outcomes *                               | Failure to conceal randomization scheme<br>Failure to achieve double-blinding<br>Failure to achieve complete follow-up  |
| 2     | Within a “head-to-head” RCT   | Identical (by definition)                                | Validated surrogate outcomes †                                | Those of level 1 PLUS<br>Validity of surrogate outcome for clinically important outcomes  |
| 2     | Across RCTs of different drugs vs. placebo  | Similar or different (in disease and risk factor status) | Clinically important outcomes or validated surrogate outcomes | Those of level 1 PLUS<br>Differences between trials in:<br>Methodologic quality (adequacy of blinding, allocation concealment, etc.)<br>Endpoint definitions<br>Compliance rates<br>Baseline risk of outcomes |
| 3     | Across subgroup analyses from RCTs of different drugs vs. placebo                           | Similar or different                                     | Clinically important outcomes or surrogate outcomes           | Those of level 1 (+ 2) PLUS<br>Multiple comparisons, post-hoc data dredging<br>Under-powered subgroups<br>Misclassification into subgroups  |
| 3     | Across RCTs of different drugs vs. placebo  | Similar or different                                     | Unvalidated surrogate outcomes                                | Surrogate outcomes may not capture all of the effects (beneficial or hazardous) of a therapeutic agent  |
| 4     | Between non-randomized studies (observational studies and administrative database research) | Similar or different                                     | Clinically important outcomes                                 | Confounding by indication, compliance, and/or calendar time<br>Unknown/unmeasured confounders<br>Measurement error<br>For outcomes research: limited databases, coding systems not suitable for research      |

\* Clinically important outcomes refer to long-term efficacy data and the particular endpoints depend on the condition being treated. For statins used to prevent/treat atherosclerotic disease, clinically important outcomes would include all-cause mortality, myocardial infarction, and stroke.

† Surrogate outcomes are considered “validated” only when the relationship between the surrogate outcome and clinically important outcomes has been firmly established in long-term RCTs.

Adapted from<sup>(19)</sup>.

classe. Por exemplo, certos bloqueadores dos canais de cálcio têm efeitos desfavoráveis sobre a mortalidade total<sup>(18)</sup> e, embora todas as estatinas actuem sobre a HMG-coenzima A redutase, elas podem exercer diferentes efeitos não lipídicos sobre o processo aterotrombótico, com potencial influência na sua eficácia clínica.

O Quadro IV apresenta uma hierarquia de níveis de evidência a usar na comparação de um fármaco com outros da mesma classe<sup>(17,19)</sup>. Aqui se podem apreciar os possíveis problemas da reclamação do efeito de classe farmacológica, quando a evidência de que dispomos apresenta qualidade e robustez variáveis.

É ainda um facto conhecido que, dentro da mesma classe, os fármacos podem ter

(association or causality), together with evidence that intervention with drugs directed at the endpoint will reduce the incidence of the disease. Moreover, there are two reasons for the importance of safety in surrogate markers: they are used as an alternative to large-scale RCTs, and a pharmacological effect that is thought to be beneficial may in fact be harmful (class 1 antiarrhythmics induce fatal arrhythmias, and beta-blockers and phosphodiesterase inhibitors increase mortality in HF)<sup>(1)</sup>.

Furthermore, in an appreciable number of cardiovascular conditions, surrogate endpoints have not been shown to be good predictors of clinical outcome (see above). Surrogate



QUADRO V

Níveis de evidência para comparação da segurança de fármacos dentro da mesma classe

| Nível | Propriedades                       | Vantagens  | Ameaças à validade  |
|-------|------------------------------------|--|---|
| 1     | RCTs                               | Único desenho que permite a detecção de efeitos adversos quando o efeito adverso é semelhante ao evento que o tratamento está a tentar impedir | Falta de poder estatístico (subpotenciado) para detectar efeitos adversos                             |
| 2     | Coorte                             | Colheita de dados prospectivos, coorte definido  | Depende criticamente do rigor do seguimento, classificação e avaliação                                |
| 3     | Caso-controlo                      | Pouco dispendioso e de execução rápida   | Viés de selecção e de lembrança; a relação temporal pode não ser clara                                |
| 4     | Vigilância pós-marketing (fase IV) | Se suficientemente ampla, pode detectar efeitos adversos raros mas importantes   | Ausência de grupo de controlo; depende criticamente do rigor do seguimento, classificação e avaliação |
| 5     | Série de casos                     | Pouco dispendioso e de execução rápida   | Amostra de pequena dimensão; viés de selecção; ausência de grupo de controlo                          |
| 6     | Casos clínicos                     | Pouco dispendioso e de execução rápida   | Amostra de pequena dimensão; viés de selecção; ausência de grupo de controlo                          |

Quadro adaptado de<sup>(19)</sup>.

TABLE V

Levels of evidence for comparing the safety of drugs within the same class

| Nível | Properties           | Advantages  | Threats to validity   |
|-------|----------------------|---|---|
| 1     | RCTs                 | Only design which permits the detection of adverse effects when the adverse effect is similar to the event which treatment is trying to prevent | Underpowered for detecting adverse effects  |
| 2     | Cohort studies       | Prospective data collection, defined cohort   | Critically depends on follow-up, classification and measurement accuracy                                  |
| 3     | Case-control studies | Cheap and fast to perform   | Selection and recall bias; temporal relationship may not be clear   |
| 3     | Phase IV studies     | If sufficiently large, can detect rare, but important, adverse effects  | No, or unmatched, control group; critically depends on follow-up, classification and measurement accuracy |
| 5     | Case series          | Cheap and fast to perform   | Small sample size; selection bias; no control group   |
| 6     | Case report(s)       | Cheap and fast to perform   | Small sample size; selection bias; no control group   |

Adapted from<sup>(19)</sup>.

perfis de segurança diferentes, um aspecto absolutamente crucial no contexto de efeito de classe. Uma abordagem similar à da eficácia torna-se imperativa no que concerne a segurança, pelo que é fundamental definir os níveis de robustez da evidência em que esta se baseia (*quadro V*)<sup>(17,19)</sup>.

Como em todas as abordagens que lançam mão de hierarquias de evidência como base de validade intrínseca das conclusões, também na

markers do not give consistent results, particularly when assessing the safety, as opposed to the efficacy, of a drug. Unless a surrogate marker has been thoroughly validated in various populations, using different drugs at different dosages and with varying follow-up periods, it does not seem scientifically defensible to make efficacy and/or safety claims for a drug without valid outcome data from RCTs.

Finally, the mere fact that a drug belongs

definição da eficácia (*quadro IV*) e da segurança (*quadro V*) dos fármacos da mesma classe se reconhece que quando não se dispõe do melhor (mais elevado) nível de evidência, pode e deve-se lançar mão de níveis mais baixos, mantendo em mente os problemas metodológicos potenciais que esta abordagem pode acarretar.

### **CONCLUSÕES: SERÁ LEGÍTIMA A RECLAMAÇÃO DE EFICÁCIA TERAPÊUTICA SEM DADOS PROVENIENTES DE ENSAIOS CLÍNICOS?**

Por tudo o que foi dito acima, parece prudente afirmar que o processo de aprovação de novos medicamentos deve basear-se primordialmente em marcadores substitutivos devidamente validados, para não se incorrer no erro de reclamação de eficácia e segurança por analogia com outros medicamentos devidamente estudados em ensaios clínicos de grandes dimensões.

Uma intervenção no MS só produz qualquer benefício directo no doente na medida em que causa ou prediz um resultado favorável (como o enfarte agudo do miocárdio (EAM) ou insuficiência cardíaca congestiva (ICC), por ex.). Os MS podem não ser válidos e um efeito adverso inesperado do medicamento pode induzir um outcome desfavorável. Os MS aparecem, em graus diversos, na cadeia causal dos eventos: o colesterol (variável bioquímica), a tensão arterial (variável fisiopatológica) e o diâmetro coronário (variável morfológica) são todos MS para EAM e ICC, mas o último é o mais importante (porque mais próximo do evento cardiovascular)<sup>(20)</sup>. O grau de persuasão sobre a validade de um MS para avaliação de benefício terapêutico está relacionado com um certo grau de plausibilidade biológica (associação ou causalidade), assim como prova que a sua modulação com fármacos diminui a incidência de doença. Finalmente, as questões de segurança com os MS assentam em dois factos: porque eles são alternativa a um RCT de grandes dimensões e o efeito farmacológico que se pensa benéfico pode ser de facto deletério (antiarrítmicos da classe I induzem arritmias letais, agonistas-beta e inibidores da fosfodiesterase aumentam a taxa de mortalidade na ICC)<sup>(1)</sup>.

to a particular therapeutic class does not support a claim of clinical equivalence, since this does not take account of its potency or other aspects that may affect its benefit, risk and cost-effectiveness<sup>(21)</sup>. The concept of a class effect is normally used only when there is little information on individual agents, particularly with regard to safety in treatment of chronic diseases, when secondary effects may not appear until years later. Drugs within the same therapeutic class should not be used interchangeably in the absence of high-quality scientific evidence comparing benefit and harm.

In conclusion, claims of equivalence between two drugs - typically when one has been thoroughly studied in large RCTs with a long follow-up and clinically relevant endpoints, and the other in small trials with a short follow-up - are legitimate and safe only when there are data from equivalence trials (or non-inferiority trials), in which there has been a direct head-to-head assessment of the two drugs in similar patient populations and using directly comparable measures of association<sup>(22)</sup>.

De resto, num número apreciável de situações cardiovasculares, os marcadores substitutivos não têm demonstrado serem bons preditores de resultados clínicos. Os MS não são consistentes no resultado do seu uso, especialmente no que concerne à segurança (*versus* eficácia) dos medicamentos. Sem um marcador substitutivo devidamente validado em diversas amostras, utilizando diversos medicamentos, com diferentes dosagens e com seguimentos variados, não parece cientificamente defensável a reclamação de eficácia/segurança de um medicamento que não possui resultados de RCTs válidos.

Finalmente, a simples pertença de uma medicamento a uma classe medicamentosa não permite reclamação de equivalência clínica, porque ignora questões como a potência ou outras acções que influenciam o benefício, risco e custo-eficácia<sup>(21)</sup>. O conceito de efeito de classe utiliza-se normalmente quando não existe grande informação sobre os agentes individuais, especialmente no que concerne à segurança em doenças crónicas, em que os efeitos secundários podem só aparecer anos mais

tarde. Medicamentos dentro da mesma classe terapêutica não devem ser permutáveis na ausência de evidência científica de alta qualidade comparando os benefícios e os danos.

Em conclusão, a reclamação de equivalência terapêutica entre dois medicamentos - tipicamente uma bem estudada em RCTs de grandes dimensões e longo termo com *endpoints* clínicos relevantes e a outra com apenas ensaios curtos de pequena dimensão - só será legítima e segura se pudermos dispor de dados de ensaios de equivalência (ou não-superioridade), em que a avaliação é feita directa e paralelamente entre as duas, nas mesmas amostras de doentes e com medidas de associação directamente comparáveis<sup>(22)</sup>.

Pedidos de separatas para:

Address for reprints:

A. VAZ CARNEIRO

Centro de Estudos de Medicina Baseada na Evidência

Faculdade de Medicina de Lisboa • 1600-035 Lisboa

E-mail: avc@fm.ul.pt

## BIBLIOGRAFIA / REFERENCES

1. Temple R. Are surrogate markers adequate to assess cardiovascular disease drugs? *JAMA* 1999;282:790-795.
2. Greenhalgh T. How to read a paper. The basics of evidence based medicine. 2nd ed. London: BMJ Publishing Group; 2001.
3. Psaty BM, Weiss NS, Furberg CD, Koepsell TD, Siscovick DS, Rosendaal FR et al. Surrogate end points, health outcomes, and the drug-approval process for the treatment of risk factors for cardiovascular disease. *JAMA* 1999;282:786-790.
4. Psaty BM, Smith NS, Siscovick DS, et al. Health outcomes associated with antihypertensive therapies used as first line agents: a systematic review and meta-analysis. *JAMA* 1997;277:739-745.
5. Soares I, Carneiro AV. Princípios de análise metodológica de ensaios terapêuticos. Parte I - Validade dos resultados. *Rev Port Cardiol* 2002;21:457-469.
6. Elwood M. Critical appraisal of epidemiological studies and clinical trials. 2nd ed. Oxford: Oxford University Press; 1998.
7. Carneiro AV. Aplicação ao doente individual dos resultados dos ensaios clínicos: regras práticas. *Rev Port Cardiol* 2003;22:259-268.
8. Carneiro AV, Costa J. Marcadores de doença cardíaca em ensaios clínicos. *Rev Port Cardiol* 2002;21:1349-1360.
9. Carneiro AV. Glossário de termos e conceitos utilizados pela cardiologia baseada na evidência científica. Parte II: Ensaios clínicos. *Rev Port Cardiol* 2001;20:203-210.
10. Jadad AR. Randomised Controlled Trials. 1st ed. London: BMJ Books; 1998.
11. Rothwell PM. External validity of randomised controlled trials: "To whom do the results of this trial apply?". *Lancet* 2005;365:82-93.
12. 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.
13. Fletcher RH, Fletcher SW. Clinical epidemiology. 4th ed. Baltimore: Lippincott Williams & Wilkins; 2005.
14. Ware JH, Antman EM. Equivalence trials. *N Engl J Med* 1997;337:1159-1161.
15. Carneiro AV. Ensaios de equivalência entre medicamentos: aspectos metodológicos. *Rev Port Cardiol* 2003;22:1125-1139.
16. McAlister FA, Laupacis A, Wells GA, Sackett DL, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. XIX. Applying clinical trial results. B. Guidelines for determining whether a drug is exerting (more than) a class effect. *JAMA* 1999;282:1371-1377.
17. Soares I, Carneiro AV. Efeitos de classe de medicamentos: definição e aplicações práticas. *Rev Port Cardiol* 2002;21:1031-1042.
18. Heidenreich PA, McDonald KM, Hastie T, Fadel B, Hagan V, Lee BK et al. Meta-analysis of trials comparing  $\beta$ -blockers, calcium antagonists, and nitrates for stable angina. *JAMA* 1999;281:1927-1936.
19. User's Guides to the Medical Literature. A Manual for Evidence-Based Clinical Practice. 1st ed. Chicago: JAMA & Archives Journals AMA; 2001.
20. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Int Med* 1996;125:605-613.
21. Furberg CD, Herrington DM, Psaty BM. Are drugs within a class interchangeable? *Lancet* 1999;354:1202-1204.
22. Kaplan NM. Should new drugs be used without outcome data? *Arch Intern Med* 2001;161:511-512.