

# Tipos de Estudos Clínicos.

## II. Estudos de *Coorte* [82]

SUSANA M. FERNANDES, ANTÓNIO VAZ CARNEIRO

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

**Rev Port Cardiol 2005 ;24 (9): 1151-1158**

### RESUMO

Os estudos de *coorte* incluem grupos de doentes que são seguidos no tempo para determinação da incidência ou da história natural da doença, assim como dos factores de risco para os vários resultados (*outcomes*).

Estes estudos permitem determinar as características das doenças e a relação de causalidade entre uma exposição e uma doença. São, no entanto, financeiramente muito pesados, porque necessitam geralmente de uma amostra de grandes dimensões.

Os estudos de *coorte* dividem-se em prospectivos (em que o investigador define a amostra e mede as variáveis preditivas antes dos resultados terem ocorrido) e retrospectivos (o investigador define a amostra e colhe dados sobre as variáveis preditivas depois dos resultados terem ocorrido). Os estudos de *coorte* estão sujeitos a uma série de viéses (erros sistemáticos), que devem ser eliminados se se quiser garantir a validade dos resultados.

Para aumentar a validade dos estudos de *coorte* é necessário que o investigador tenha em atenção a exposição, o *outcome*, a escolha da amostra e a forma como é feita a análise estatística (rácios ou diferenças de risco, riscos absolutos e relativos, etc.).

### Palavras-Chave

Estudos clínicos; Estudos de *coorte*; Estudos longitudinais, prospectivos e retrospectivos; Epidemiologia clínica; Cardiologia baseada na evidência

### ABSTRACT

#### Types of clinical studies.

#### II. Cohort studies

Cohort studies include groups of patients that are followed over time to determine the incidence or natural history of a disease, together with the risk factors for specific outcomes. These studies can be used to determine disease characteristics as well as the causal relationship between exposure and disease. They are, however, very expensive, because they need large patient samples.

Cohort studies can be classified as prospective (the investigator defines the sample and measures predictive variables before outcomes arise) or retrospective (the investigator defines the sample and measures predictive variables after outcomes arise). Cohort studies are susceptible to various types of bias (systematic errors) that must be eliminated in order to ensure the validity of the results.

To increase the validity of cohort studies the investigator has to consider exposure, outcome, sample selection and the statistical analysis (risk ratios and/or differences, absolute and relative risks, etc.).

### Key words

Clinical studies; Cohort study; Longitudinal, prospective and retrospective studies; Clinical epidemiology; Evidence-based cardiology

## INTRODUÇÃO

Este é o terceiro de uma série de artigos em que procuramos – como já foi sublinhado previamente<sup>(1)</sup> – explicitar a construção e a análise dos vários tipos de estudos, discutindo cada um separadamente. No número anterior da Revista foi apresentado e discutido o formato dos estudos caso-controlo. Neste artigo focar-nos-emos sobre os estudos de *coorte*, que são os estudos observacionais que mais se aproximam dos ensaios clínicos aleatorizados e controlados.

Uma *coorte* é um grupo de indivíduos que têm qualquer coisa em comum entre eles e que permanecem como um grupo durante um determinado período de tempo<sup>(2)</sup>. Os estudos de *coorte* têm geralmente dois objectivos principais: descrever a incidência de *outcomes* (resultados) (estudos descritivos), ou analisar as associações entre esses *outcomes* e factores preditores dos mesmos (estudos analíticos)<sup>(3)</sup>. Não são porém estudos que determinem causalidade, visto serem apenas observacionais, permitindo porém estabelecer a existência de associações entre exposição e resultado<sup>(4)</sup>.

Para avaliar o significado dos dados obtidos através de um estudo de *coorte*, de modo a melhorar a sua validade, é importante ter em conta várias questões<sup>(5)</sup>:

1. Qual a hipótese em estudo?
2. Qual o desenho do estudo?
3. Qual a população em estudo?
4. Qual a exposição?
5. Qual a doença a que se pretende associar a exposição?
6. Como foi efectuado o seguimento?
7. Como foi efectuada a análise estatística?

Neste trabalho pretendemos apontar os aspectos importantes relacionados com estas questões, incidindo sobre a informação que um estudo de *coorte* deve fornecer. Iremos ainda analisar as vantagens e desvantagens dos estudos de *coorte*.

## DESENHO METODOLÓGICO

### 1. A hipótese

Em primeiro lugar é necessário compreender-se o tipo de questão que foi colocada. Os estudos de *coorte* são particularmente úteis para determinar a relação entre uma exposição e um *outcome* – principalmente quando as exposições

## INTRODUCTION

This is the third in a series of articles in which, as outlined previously<sup>(1)</sup>, we set out to explain the design and analysis of various types of study, describing each one separately. In the last issue of the Journal, we presented and discussed case-control studies. In this article we focus on cohort studies, which of all observational studies are the closest to randomized clinical trials.

A cohort is a group of individuals who have something in common and who remain part of the group for a particular period of time<sup>(2)</sup>. Cohort studies usually have one of two main objectives: to determine the incidence of outcomes (descriptive studies), or to analyze the associations between these outcomes and their predictive factors (analytical studies)<sup>(3)</sup>. However, being merely observational they cannot determine causality, although they can establish associations between exposure and outcome<sup>(4)</sup>.

To evaluate the significance of data obtained from a cohort study and to increase its validity, certain questions should be borne in mind<sup>(5)</sup>:

1. What is the study hypothesis?
2. What is the study design?
3. What is the study population?
4. What is the exposure?
5. What disease is associated with the exposure?
6. How was follow-up effected?
7. How was the statistical analysis performed?

In this article we shall examine the most important aspects of these questions, with particular emphasis on the information that cohort studies can provide. We shall also describe their advantages and disadvantages.

## METHODOLOGICAL DESIGN

### 1. The hypothesis

It is first necessary to understand what type of question is being posed. Cohort studies are particularly useful for determining the relationship between an exposure and an outcome, especially when the exposure is rare, no longer occurs, or cannot for ethical reasons be deliberately inflicted on subjects<sup>(5)</sup>. An example would be the information provided by a cohort study on the risk of coronary disease after radiotherapy: it would not be ethical to expose

são raras, já não ocorrem, ou não é ético expor propositadamente o indivíduo<sup>(5)</sup>. Por exemplo, um estudo de *coorte* pode dar-nos informação sobre o risco de doença coronária após radioterapia: não é ético expor indivíduos saudáveis a radioterapia, mas pode-se escolher uma *coorte* de doentes com neoplasia da mama que tenha feito radioterapia, comparando-a com uma que não a tenha feito, seguir prospectivamente os dois grupos ao longo do tempo, observando então o aparecimento de doença coronária<sup>(6)</sup>.

## 2. Desenho

O desenho destes estudos envolve a escolha de dois grupos de indivíduos – duas *coortes* – uma de expostos e outra de não expostos a um determinado factor (eventualmente) causal, quantificando-se os que desenvolvem doença. De uma forma geral estes estudos podem ter duas abordagens: prospectiva e retrospectiva (*Fig. 1*).

Habitualmente os estudos são prospectivos, estudando-se as *coortes* ao longo do tempo, esperando que ocorra o *outcome* previamente definido (*Quadro 1*). Este tipo de estudos é também designado como longitudinal<sup>(5)</sup>.

Quando a exposição é frequente e o estudo é prospectivo, pode escolher-se uma população estável, para que não haja perdas de seguimento, dividindo-a posteriormente pelos dois

healthy individuals to radiotherapy, but a cohort of breast cancer patients who have undergone radiotherapy can be selected and compared with another who have not, and the two groups can be followed prospectively over time and the appearance of coronary disease monitored<sup>(6)</sup>.

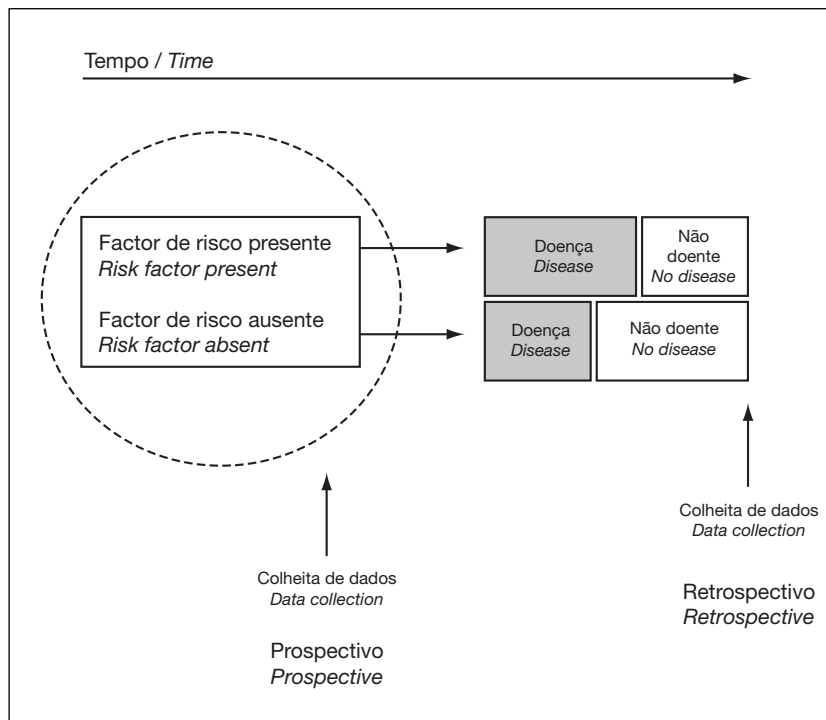
## 2. Design

The design of these studies involves the selection of two groups of individuals (cohorts), one that has and the other that has not been exposed to a (possibly causal) factor, and those that develop the disease are then quantified. Cohort studies may be prospective or retrospective (see *Fig. 1*).

Most studies are prospective, studying the cohorts over time and waiting for the previously defined outcome (*Table 1*). Such studies are also termed longitudinal<sup>(5)</sup>.

When the exposure is frequent and the study is prospective, a stable population can be selected, such as a group of nurses or the population of a city, and subsequently divided into the two groups, in order to minimize losses to follow-up.

In retrospective cohort studies, both the exposure and the outcomes occurred in the past. In this case the investigator uses clinical records or interviews to determine the presence or otherwise of exposure and the consequent outcome.



*Fig. 1* Esquema do desenho dos estudos de *coorte* (adaptado de<sup>(3)</sup>).

*Fig. 1* Diagram of the design of cohort studies (adapted from<sup>(3)</sup>).

*Quadro I*  
**Comparação entre estudos retrospectivos e prospectivos**

Atributos	Retrospectivos	Prospectivos
Informação	Menos completa	Mais completa e correcta
Exposições descontinuadas	Útil	Não útil
Novas exposições emergentes	Não útil	Útil
Custo	Baixo	Alto
Tempo do estudo	Curto	Longo

Adaptado de<sup>(6)</sup>.

grupos (por exemplo uma população de enfermeiras, ou uma população da mesma cidade).

Nos estudos de *coorte* retrospectivos, tanto a exposição como os *outcomes* já ocorreram no passado. Neste caso, o investigador serve-se de registos clínicos ou entrevistas para determinar a existência de uma exposição e a presença ou não do *outcome* consequente.

### 3. População

São vários os factores que influenciam a selecção dos indivíduos: por exemplo, a exposição em investigação, ou a frequência da exposição na população, ou a acessibilidade aos indivíduos, ou a probabilidade dos indivíduos continuarem no estudo<sup>(7)</sup>.

Para que o estudo seja significativo é essencial que ambos os grupos dos pacientes expostos e não expostos não apresentem o *outcome* na altura da exposição, e que o grupo dos não expostos tenha características em tudo semelhantes ao dos expostos.

### 4. Exposição

São várias as características da exposição que determinam a escolha da *coorte*: rara ou frequente, aguda ou crónica, determinados níveis de exposição. Por exemplo, é possível na avaliação de uma exposição rara incluírem-se todos indivíduos expostos. Além disso, se a exposição incluir a existência de várias doses, pode também considerar-se a hipótese de determinar uma curva de dose-resposta, por exemplo a relação entre hábitos tabágicos ligeiros, moderados e intensos e a probabilidade de doença cardiovascular.

### 5. Outcome

Aqui é importante que o investigador defina de uma forma clara o *outcome*, com especial ên-

*Table I*  
**Comparison between retrospective and prospective studies**

Atributes	Retrospective	Prospective
Information	Less complete	More complete and accurate
Discontinuous exposures	Useful	Not useful
Emerging exposures	Not useful	Useful
Cost	Low	High
Study period	Short	Long

Adapted from<sup>(6)</sup>.

### 3. Population

There are various factors that influence the selection of the sample, including the exposure under study, the frequency of exposure in the population, how accessible the subjects are, and the likelihood that they will remain in the study<sup>(7)</sup>.

If the study is to be significant, both groups – those that are exposed and those that are not – must not have the outcome at the time of exposure, and the non-exposed group must have exactly the same characteristics as the exposed group.

### 4. Exposure

Different aspects of the exposure influence the selection of the cohort: it can be common or uncommon, acute or chronic, and at different levels. For a rare exposure, for instance, all exposed individuals can be included; for different levels of exposure, it may be possible to construct a dose-response curve, such as the relation between light, moderate and heavy smoking and the risk of cardiovascular disease.

### 5. Outcomes

It is important to define the outcome clearly, particularly with regard to diagnosis. In a retrospective study, the presence or otherwise of the outcome is determined based on clinical records or interviews. If the conclusions are to be valid, both groups must have undergone the same diagnostic tests, and the person responsible for the diagnosis must be blinded to the groups to which the subjects belong<sup>(8)</sup>.

### 6. Follow-up

In prospective cohort studies the follow-up period is usually long, since subjects must be followed for sufficient time for the outcome to

fase sobre o método de diagnóstico. Num estudo retrospectivo recorre-se a registos ou entrevistas para determinar a presença ou não do *outcome*. Para que as conclusões sejam válidas, é necessário que ambos os grupos tenham sido sujeitos aos mesmos testes diagnósticos, sendo também importante verificar se quem fez o diagnóstico permanece em ocultação para o grupo a que pertencem os indivíduos<sup>(6)</sup>.

## 6. Seguimento

Nos estudos de *coorte* prospectivos o seguimento é geralmente longo, já que para que o *outcome* possa surgir, é necessário que os indivíduos sejam seguidos durante tempo suficiente, o que por vezes significa a passagem de vários anos. Estes tempos de seguimento tornam por vezes difícil manter registos de todos os doentes, o que obriga o investigador a desenvolver medidas para manter os indivíduos no estudo: aceitam-se pequenas perdas de seguimento, porém grandes perdas podem invalidar todo o estudo.

## 7. Análise estatística

Os resultados de um estudo de *coorte* podem ser sumariados num formato semelhante ao *Quadro II*<sup>(6)</sup>:

*Quadro II*

### Resultados de um estudo de coorte

	Expostos	Não expostos	Total
<i>Outcome</i> presente	A	B	A+B
<i>Outcome</i> ausente	C	D	C+D
Total	A+C	B+D	A+B+C+D

Neste Quadro os grupos A e C correspondem aos que foram expostos, sendo B e D os que não foram expostos. Os grupos A e B atingiram o *outcome*, os grupos C e D não o atingiram.

De entre os indivíduos expostos, o risco de atingir o *outcome* pode então ser definido como:

$$R(\text{expostos}) = \frac{\text{Expostos com outcome}}{\text{Todos os expostos}} = A / (A+C)$$

O risco pode então variar entre zero (nenhum dos expostos teve o *outcome*) e um (todos os expostos atingiram *outcome*):

$$R(\text{não expostos}) = \frac{\text{Não expostos com outcome}}{\text{Todos os não expostos}} = B / (B+D)$$

appear, which may be years. This can make it difficult to maintain records of all the patients, which means the investigators must find ways to retain the subjects in the study. Small losses to follow-up are acceptable, but large ones may invalidate the entire study.

## 7. Statistical analysis

The results of a cohort study can be summarized as in *Table II*<sup>(6)</sup>:

*Table II*

### Results of a cohort study

	Exposed	Non-exposed	Total
Outcome present	A	B	A+B
Outcome absent	C	D	C+D
Total	A+C	B+D	A+B+C+D

In this table, groups A and C are those who have been exposed, and B and D those who have not. Groups A and B have had the outcome, while C and D have not.

Among the exposed subjects, the risk of having the outcome can be defined as

$$R(\text{exposed}) = \frac{\text{Exposed with the outcome}}{\text{All those exposed}} = A / (A+C)$$

The risk may thus range between 0 (none of those exposed had the outcome) and 1 (all of those exposed had the outcome).

$$R(\text{non-exposed}) = \frac{\text{Non-exposed with the outcome}}{\text{All those non-exposed}} = B / (B+D)$$

To interpret these risks, they must be related by calculating the relative risk.

$$RR(\text{relative risk}) = \frac{R(\text{exposed})}{R(\text{non-exposed})} = [A / (A+C)] / [B / (B+D)] = A(B + D) / B(A + C)$$

If the relative risk equals 1, the risk is the same in both groups. If it is greater than 1 the exposure is harmful, while if it is less, the exposure is beneficial. Relative risk is thus a measure of the strength of association between the exposure and the outcome: the further it is from unity, the stronger the association. The accuracy of this figure can be improved by calculating confidence intervals; if they do not include 1, the result is statistically significant and the exposure can be associated with the outcome.

Para conseguirmos interpretar estes riscos é necessário relacioná-los entre si – é o risco relativo.

$$RR = R(\text{expostos}) / R(\text{não expostos}) \\ = [A / (A+C)] / [B / (B+D)] = A(B + D) / B(A + C)$$

Se o risco relativo for igual a 1, então o risco é igual nos dois grupos. Se o risco for maior do que 1 a exposição é deletéria, se for menor que 1 é benéfica. O risco relativo torna-se assim uma medida de força de associação entre a exposição e *outcome*: quanto mais distante for da unidade, mais forte é a associação. Para acrescentar precisão a este valor calculam-se os intervalos de confiança. Se não incluírem 1, então existe significância estatística e é possível associar a exposição ao *outcome*.

Uma outra medida útil é o risco absoluto que se define por:

$$RA = \text{Risco expostos} - \text{Risco não expostos}$$

Esta medida dá-nos o excesso de risco em que o indivíduo incorre por ter estado exposto. A percentagem de risco absoluto define-se assim:

$$PRA = [RA / R(\text{expostos})] \times 100$$

É uma outra medida, tipicamente usada como indicador do impacto público de determinada exposição.

### VANTAGENS E DESVANTAGENS DOS ESTUDOS DE COORTE

Como em todos os outros tipos de estudo, o estudo de *coorte* apresenta vantagens e desvantagens (*Quadro III*). Este desenho de estudo é excelente para estabelecer relação causa-efeito.

#### Quadro III

#### Vantagens e desvantagens dos estudos de coorte

Vantagens	Desvantagens
Cálculo directo do risco relativo	Tempo
Informação sobre a incidência da doença	Requer grupos muito grandes
Relação temporal clara entre exposição e doença	Caro
Eficiente para estudo de exposições raras	Não é eficiente para estudo de doenças raras
Informação sobre múltiplas exposições	Seguimento difícil
Informação sobre múltiplos <i>outcomes</i> relacionados com apenas uma exposição	Modificação ao longo do tempo dos métodos de diagnóstico pode enviesar o estudo
Viés mínimos	
Estudo observacional que melhor estabelece relação causa efeito	

Adaptado de<sup>(6)</sup>.

Another useful measure is absolute risk, defined as

$$AR = \text{Risk of those exposed} - \text{Risk of those non-exposed}$$

This measure shows the added risk incurred by exposure. The percentage absolute risk is calculated by

$$PAR = [AR / R(\text{exposed})] \times 100$$

This measure is frequently used as an indicator of the public impact of a particular exposure.

### ADVANTAGES AND DISADVANTAGES OF COHORT STUDIES

Like all types of study, cohort studies have advantages and disadvantages (*Table III*). Their design is excellent for establishing causal relationships, and they can also provide data on disease incidence and prognosis, the association between various exposures and outcomes, and the consequences of rare exposures.

They do, however, have several limitations. The study sample needs to be large to obtain statistically significant results; for a rare outcome, the number of subjects required is so great that a cohort study may be impractical. They also usually require a long follow-up period, which increases their cost. Further, the investigator must bear in mind that new diagnostic methods may be developed during the follow-up period. Some of these disadvantages can be overcome in retrospective studies. However, what these gain in reduced costs is lost in reduced validity, since their dependence on memory or records means that retrospective studies are often subject to bias.

Table III

**Advantages and disadvantages of cohort studies**

Advantages	Disadvantages
Direct calculation of relative risk	Time
Information on disease incidence	Require large samples
Clear temporal relation between exposure and disease	Expensive
Efficient for studying rare exposures	Inefficient for studying rare diseases
Information on multiple exposures	Difficult follow-up
Information on multiple outcomes related to a single exposure	Changes over time in diagnostic methods can lead to bias
Minimal bias	
The best observational study for establishing causal relationships	

Adapted from <sup>(9)</sup>.

Se forem efectuados para isso, podem fornecer dados sobre a incidência e prognóstico de doenças, associação entre várias exposições e *outcomes*, ou consequências de exposições raras.

Apresentam porém várias limitações: por um lado, necessitam de um grande número de indivíduos para terem resultados estatisticamente significativos (na realidade, quando um *outcome* é raro, a quantidade de indivíduos necessária é tão grande que se torna incomportável a sua realização). Por outro, implicam geralmente um tempo de seguimento prolongado, o que também aumenta o custo do estudo. Finalmente, o investigador tem também que contar com a possibilidade de surgirem entretanto novos métodos de diagnóstico, já que o tempo de seguimento é muito prolongado. Algumas destas desvantagens são colmatadas pelos estudos retrospectivos. Porém, o que estes ganham em diminuição dos custos perdem em validade, já que apresentam habitualmente viés de memória ou nos registos.

**CONCLUSÕES**

Como vimos ao longo deste artigo, os estudos de *coorte* são uma boa metodologia para avaliar prevalência de doença, assim como associações entre uma exposição e um *outcome*, sendo particularmente útil para exposições raras. Por outro lado são estudos caros e pouco úteis para *outcomes* raros.

Na avaliação este tipo de estudos é importante que seja observada a forma como foi feita a escolha dos grupos de exposição e não exposição, assim como deverá haver garantia de que a análise estatística respectiva é cuidadosa e rigorosa.

**CONCLUSIONS**

As we have seen in this article, cohort studies are a good method for studying disease prevalence and associations between exposure and disease, being particularly useful for rare exposures. They are, however, expensive and of little use for rare outcomes.

When evaluating cohort studies, it is important to check how the samples of exposed and non-exposed subjects were selected and that the statistical analysis was careful and rigorous.

Pedidos de separatas para:  
 Address for reprints:

ANTÓNIO VAZ CARNEIRO  
 Centro de Estudos de Medicina Baseada na Evidência  
 Faculdade de Medicina de Lisboa  
 1649-028 LISBOA, PORTUGAL  
 e-mail: avc@fm.ul.pt

## BIBLIOGRAFIA / REFERENCES

1. Carneiro AV. Tipos de estudos clínicos. Introdução. Rev Port Cardiol 2005;24:897-903.
2. Dawson B, Trapp RG. Basic & clinical biostatistics. 3rd ed. New York: Lange Medical Books/McGraw-Hill, 2001.
3. Hulley SB, Cummings SR, Browner WS, Grady D, Hearst N, Newman TB. Designing clinical research. 2nd ed. Philadelphia: Lippincott Williams and Wilkins, 2001.
4. Grimes DA, Schulz KF. Cohort studies: marching towards outcomes. Lancet 2002;359:341-5.
5. Greenberg RS, Daniels SR, Flanders WD, Eley JW, Boring JR. Medical epidemiology. 3rd ed. New York: Lange Medical Books/McGraw-Hill, 2001.
6. Darby S, McGale P, Peto R, Granath F, Hall P, Ekblom A. Mortality from cardiovascular disease more than 10 years after radiotherapy for breast cancer: nationwide cohort study of 90000 Swedish women. BMJ 2003;326:256-7.
7. Armitage P, Berry G, Matthews JNS. Statistical methods in medical research. 4th ed. Oxford: Blackwell Science, 2002.
8. Lang TA, Secic M. How to report statistics in medicine. 1st ed. Philadelphia: ACP, 1997.