JHGD
www. jhgd.com.br
ORIGINAL ARTICLE

Research methodology topics: Cohort studies or prospective and retrospective cohort studies

Luís Marcelo Aranha Camargo[1,2,3,4,5], Romeu Paulo Martins Silva[5,6], Dionatas Ulises de Oliveira Meneguetti[5,7]

Open acess

[1]Institute of Biomedical Sciences, University of São Paulo, Monte Negro, Rondônia, Brazil

[2]Department of Medicine, São Lucas University Center, Porto Velho, Rondônia, Brazil.

[3]Center of Tropical Medicine of Rondônia, Porto Velho, Rondônia, Brazil.

[4]National Institute of Science and Technology EpiAmo / Rondônia.

[5]Stricto Sensu Graduate Program in Health Science in the Western Amazon, Federal University of Acre, Rio Branco, Acre, Brazil.

[6]Center for Health and Sports Sciences, Federal University of Acre, Rio Branco, Acre, Brazil.

[7]School of Application, Federal University of Acre, Rio Branco, Acre, Brazil.

Corresponding author
dionatas@icbusp.org

Manuscript received: September 2018
Manuscript accepted: August 2019
Version of record online: October 2019

Abstract

In health sciences, the epidemiological method can be divided into descriptive and analytical epidemiology and the latter being divided into observational (cross-sectional study, case-control study and cohort study) and experimental studies. Cohort studies may be retrospective or prospective, and both assume that the researcher will follow a population over time to seek a possible association between exposure (s) and outcome(s). These types of studies have as advantages the possibility of measuring several exposure factors and outcomes, both primary and secondary, for both relatively frequent outcomes and rare exposure factors. However, they are often long and therefore expensive studies. They have as main biases those of selection, memory and information. These are studies that may point to statistical associations between exposure and outcome that need other models to prove the casualty of these associations.

Keywords: cohort, longitudinal study, follow-up.

Suggested citation: Camargo LMA, Silva RPM, Meneguetti DUO. Research methodology topics: Cohort studies or prospective and retrospective cohort studies. *J Hum Growth Dev. 2019; 29(3):433-436.* DOI: https://doi.org/10.7322/jhgd.v29.9543

Authors summary

Why was this study done?
The present study was conducted to show the advantages and biases of a prospective and retrospective cohort study, demonstrating its application and in which situations it is indicated.

What did the researchers do and find?
It was found that the cohort study may be retrospective and prospective. In the retrospective study the researcher collects previous information on exposure factors and over time the in individuals. In the prospective study, the researcher is present at the exposure of one or more factors and followed for a period of time to observe one or more outcomes.

What do these findings mean?
That the cohort study, even with some biases, is a method that can estimate the incidence of an outcome (or more) exposed to one or more factors, and verifies whether there is a statistical association between exposure and outcome, either primary and secondary.

INTRODUCTION

The epidemiological method, in the area of Health Sciences, can be divided into two main areas: descriptive and analytical epidemiology. In the first part, indicators of morbidity, mortality, demographic, socioeconomic, quality of access health services and quality of life are used, among others. The second part, on other hand, uses data from descriptive epidemiology to analyze and seek explanations and/or associations for the descriptive data[1-3]. These (analytical) studies, on other hand, are divided into observational (cross-sectional study[4], case-control study and cohort study) and experimental studies, best known as clinical trials[1-3]. They all have different advantages, disadvantages, costs, runtime and accuracy. While observational studies raise more hypothesis rather than showing causal association, experimental studies, considered the gold standard of epidemiological studies, are expensive, time-consuming but more accurate and may show a causal relationship between exposure and outcome.

COORT STUDY

These studies (also called longitudinal or "follow-up") assume that the researcher will follow a population over time to seek a possible (at least statistical) association between exposure and outcome. Roughly this study can be divided into 2 subtypes: a retrospective cohort study and a prospective cohort study (Figure 1).

In the prospective study the researcher is present at the time of exposure of one or more factors and accompanies for a period of time to observe one or more outcomes. In the case of the retrospective study the researcher can collect previous information on the exposure factor (s) (hence the retrospective term) and accompany individuals for a period of time (the cohort).

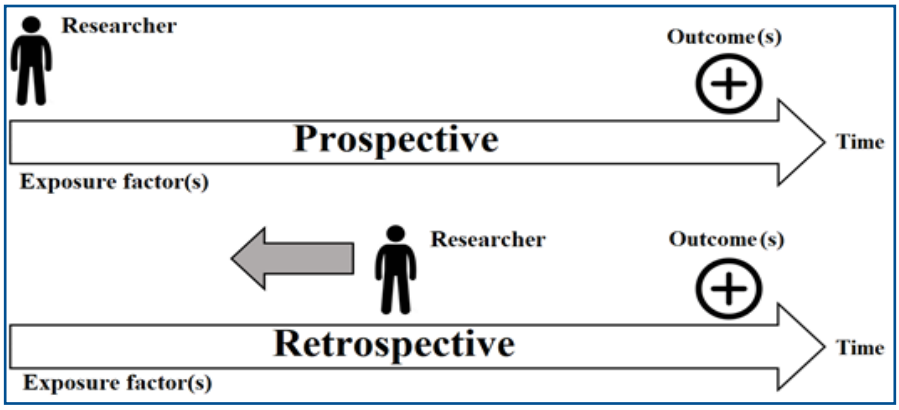

Figure 1: Prospective and retrospective cohort study.

After some time of monitoring (months, years or decades) the researcher can relate the exposure to the outcome(s) using the Relative Risk (RR), which is nothing more than the incidence of the outcome (s) in the cohort. At the end of the study the investigator analyzes the incidence of the outcome in the group of exposed and non-exposed to the risk factor (s) in a contingency table (Table 1).

The RR will be calculated using the following formula:

This formula, in summary, represents the ratio of the Incidence between exposed (I_e) and Incidence between unexposed (I_u), therefore $RR = I_e / I_u$.

The calculation of the sample size for a coort study is given by the following formula:

$$n = \frac{N \cdot Z^2 \cdot p \cdot (1 - p)}{Z^2 \cdot p \cdot (1 - p) + e^2 \cdot (N - 1)}$$

(n - Calculated sample; N - Population; Z - Normal standardized variable associated with the confidence level; p - True probability of the event; e - Sample error)

Table 1: Incidence of outcome versus exposure

Exposure/Outcome factor	Outcome +	Outcome -	Total
Exposure +	A	B	A+B
Exposure –	C	D	C+D
Total	A+C	B+D	A+B+C+D

Both RR calculation and sample size can be performed using open access programs such as EpiInfo[5] and OpenEpi[6]. There are also other online platforms that perform these types of analyzes, and allows the calculation of sample size[7].

In addition to estimating the incidence of an outcome (or more) exposed to one or more exposure factors, one can verify whether there is a statistical association between exposure and outcome. The chi-square test can be used to verify the 95% confidence interval. If the number 1 is excluded from the confidence interval and / or the statistical significance (p) calculation is less than 5% (0.05), the statistical association can be validated (although it can not be assured that there is a causal relationship).

As an example of a prospective cohort study, sets up the following situation: a cohort of 3,400 people followed for 10 years, of whom 1,600 were exposed to a chemical and another 1,800 who were not exposed. As the outcome, the incidence of cases of myeloid leukemia was measured. In the exposed group there were 100 cases and in the group without product exposure (25) cases in 10 years (prospective cohort study). In order to calculate the RR using the formula above, we elaborate the one of contingency below (Table 2).

Table 2: Example of the contingency table related to the above example.

Exposure / Outcome Factor	Leukemia +	Leukemia -	Total
Product Exposure +	100	1.500	1.600
Product Exposure –	25	1.775	1.800
Total	125	3.275	3.400

Using the formula for calculating RR above:

The RR =4.5 (CI 95% 2.918- 6.939)

* Interval Confidence 95% (95% IC) and p <0.05 (statistically significant).

With this data we can affirm that the occurrence of new cases of leukemia in the group exposed to the chemical has an incidence (Ie) of 62.55 new cases / 1,000 people, which in the group not exposed to incidence (Iu) is 14.1 /1,000 people and the statistically significant risk of leukemia in the exposed group is 4.5 higher than in the non-exposed group (RR) with high statistical significance (p = 0.001). Since p <0.05 and number 1 is not in the confidence interval, I accept the hypothesis of statistical association between outcome exposition.

In addition to the calculation of RR and incidence between exposed and non-exposed, this study model allows to calculate (Table 3).

Table 3: Other possible measures to be obtained using a cohort study.

Total Cumulative Incidence (TCI)	TCI= [(a+c)/a+b+c+d]. 10^n
Attributable Risk (AR)	AR= I_e- I_o for 10^n It estimates the excess risk of the disease in the population exposed to the risk factor
Percentage of Attributable Risk (PAR)	$PAR=\frac{(RR-1)}{RR}.100$ Estimates the percentage of disease attributable to exposure
Risk Attributable to the Population (RAP)	RAP= (IAT- I_o) .10n Estimates the excess risk of disease in the population attributable to exposure
Risk Attributable to the Population in % RAP%	$RAP\%=\frac{(IAT- I_o)}{IAT}.100$ Percentage of disease in population attributable to exposure

CONCLUSION

Cohort studies present as advantages the ability of measuring several exposure factors and outcomes (primary and secondary). They can be applied to relatively frequent outcomes and rare exposure factors. In general, these studies are relatively expensive due to the cohort follow-up time. The errors of information and memory are the major biases of these studies. Often, in the presence of a statistical association between exposure and outcome, the Clinical Trial (experimental) is sought to confirm the hypothesis.

Perhaps the most successful example of Cohort Study (prospective) is the Framingham Study that has been going on for over 60 years. Much of the information on risk factors (exposure) and cardiovascular diseases (outcomes) derive from this study.

Acknowledgments

To Sergio de Almeida Basano and Juliana de Souza Almeida A. Camargo for the critical review of the article

REFERENCES

1. Gordis L. Epidemiology. 5th Ed. Saunders Elsevier. 2013.
2. Galvão TF, Pereira MG, Silva MT. Saúde baseada em evidências. Guanabara Koogan, 2016.
3. Merril RM. Introduction to epidemiology. 5th ed. Jones & Bartlett Publishers, 2010.
4. Zangirolami-Raimundo J, Echeimberg JO, Leone C. Research methodology topics: Cross-sectional studies. J Hum Growth Dev. 2018;28(3):356-60. DOI: http://dx.doi.org/10.7322/jhgd.152198
5. Centers for Disease Control and Prevention (CDC). Epi InfoTM. [cited 2018 Dec 12] Available from: https://www.cdc.gov/epiinfo/index.html
6. OpenEpi. Estatísticas epidemiológicas de código aberto para a Saúde Pública. OpenEpi. [cited 2018 Dec 12] Available from: https://www.openepi.com/Menu/OE_Menu.htm
7. Santos GEO. Cálculo amostral. [cited 2018 Dec 14] Available from: https://praticaclinica.com.br/anexos/ccolaborativa-calculo-amostral/ccolaborativa-calculo-amostral.php

Resumo

Na área de ciências da saúde, o método epidemiológico, pode ser dividido em epidemiologia descritiva e a analítica, essa última se divide em observacional (estudo de corte transversal, estudo caso-controle e estudo de coorte/cohorte) e experimentais. Os estudos de coorte ou cohorte, podem ser retrospectivos ou prospectivos, e ambos partem do pressuposto que o pesquisador irá acompanhar uma população ao longo do tempo para buscar possível associação entre exposição e desfecho. Esses tipos de estudos apresentam como vantagens a possibilidade de se mensurar vários fatores de exposição e desfechos, tanto primários como secundários, aplicam-se tanto para desfechos relativamente frequentes e fatores de exposição raros. Porém, muitas vezes são estudos prolongados e, portanto, caros. Têm como principais viéses os de seleção, memória e informação. São estudos que podem apontar para associações estatísticas entre exposição e desfecho que necessitam de outros modelos para se comprovar há casualidade destas associações.

Palavras-chave: coorte/cohorte, estudo longitudinais, follow-up.