STATISTICS & MEDICINE: THE EXACT LINK

BERGHICHE A

Laboratory science and technic of living, Institute of Agronomic and Veterinary Sciences, University of Mohamed Cherif Messaâdia, Souk Ahras, Algeria.

**Corresponding author: amine_berghiche@yahoo.com*

Introduction

Statistics is the set of techniques and procedures used for the collection, description; analysis and interpretation of data, in the medicines' there are many questions that can only be answered by statistics. (Everitt & Skrondal., 2002). In the field of medical science, clinicians are always required to understand and make judgements. By judging, we mean making decisions about causal relationships, diagnoses or therapeutic decisions, and making decisions based on knowledge of the facts and the underlying mechanisms. In effect, decision-making is the impact of medical research, and at this level we have defined epidemiology perfectly (Zhou et al., 2009). According to (Wunsh, 2010) In the social sciences, causal histories are "modelled" (theoretical approach) and one sees to what extent the model "fits" with the known facts, but in the biomedical sciences the approach is different; It starts with observation and ends with the "model". It is called inductive. Scientific knowledge has accumulated more through "reason and intuition" than through empirical observation. Indeed, observation alone is not scientific if it does not lead to a certain prediction obtained by a certain inductive logic.

However, inductive logic is not everything because it often leads to error. On the other hand, deductive logic can be complete. For this, it must be linked to its natural environment. In fact, the history of the disease is derived by abstraction and generalisation from numerous concrete cases. This results in a picture of the disease with its common forms, clinical forms and complications. In order to do this, a maximum amount of data must be collected, which makes it possible to describe the picture of the disease and its spread. The basis of causal research is the emphasis on the temporal order: "the effect never precedes the cause" (Davis, 1985). In this regard, the mathematical method is used to confirm or reject certain equivocal points. Statistical analysis is of considerable value in causal research. Its purpose is to ascertain the degree of connection between the effect and the cause, and if so, what the probability of this connection is (Bagozzi & Yi, 1988). Thus, biostatistics does not define the cause, but that there are criteria for the causal link that can be formulated in terms of probability theory. Statistical analysis does not deal with a case but with classes of facts or variables. It therefore assumes that the work of description has already been done. (Holland, 1986).

Epidemiology, like all sciences, aims to contribute to the establishment of scientific truth through the accurate representation of health phenomena. (Polgar & Thomas, 2011). In particular by ensuring the reduction and elimination of accidental and systematic errors (bias) in the presentation, processing and analysis of data with a view to their unbiased interpretation, processing and analysis of data with a view to their unbiased interpretation. (Anderson et al., 2009).

Epidemiology should enable scientific reasoning by strengthening the criteria for clinical and community health judgment. In practical terms, the aims of epidemiology are to contribute to a better definition of diseases and to help classify them; to identify the extent of health problems and diseases and describe their characteristics; to contribute to the identification of causal relationships; to contribute to the evaluation of the effectiveness of health programmes (primary prevention, treatment, care, modification of health services, etc.); and to contribute to a better understanding of the impact of health services on the environment. (Bonita et al., 2006). Treatment, care, behavioural change, rehabilitation) and to conduct epidemiological surveillance, to study the long-term evolution of health phenomena under changing natural conditions. (Hawley & Altizer, 2011). The evolution of epidemiology requires a classification of epidemiological studies according to the purpose of the study. A general distinction can be made between observational epidemiological studies and interventional and/or experimental epidemiological studies, intervention and/or experimental studies. (Röhrig, et al., 2009).

Descriptive studies are studies limited to the description of health phenomena and are fact-finding studies whose objective is to develop a research hypothesis, that are not research-oriented are surveillance studies surveillance studies, known as epidemiological surveillance. The second group of studies are analytical studies, which are start with a hypothesis and aim to test the validity of this hypothesis; this method consists of testing the relationship between a health condition and certain variables. (Kim et al., 2017).

Epidemiological surveillance is not about knowing the health phenomenon. In this case the cause, the risk and the disease are known. The purpose of epidemiological surveillance is individual prevention. In this context, it is a continuous monitoring of individuals exposed to a particular risk, where all the data on risk factors and their possible effects are collected (Declich & Carter, 1994). The contribution of informatic tools is considerable; databases have been set up such as EUROCAT (congenital malformations), EUROTRANSPLANT (database for organ donor identification), etc. (Haux, 2006 ; Pratschke et al., 2019). In an epidemiological survey, the choice of study subjects or a representative sample from the general population necessarily requires the systematic use of a random sampling method. (Martínez-Mesa et al., 2016).

A random sample is a probability method that is defined by the fact that that all the subjects that make up a given parent population must have the same probability of being the same probability of being part of the sample. (Kothari, 2004) In technical terms, random surveys are based on the principle of drawing lots from the population, based on the principle of drawing lots from a number of subjects who can be considered who can be considered representative of the parent population and the draws must be carried out anonymously. (Trochim & Donnelly, 2001). In the random sampling method, the representativeness of the sample is a fundamental rule. fundamental rule; Despite the strict observance of the random draw, the possibility of error is nevertheless envisaged in random surveys. However, despite the strict observance of the random draw, the possibility of error and of a non-representative sample is envisaged in random surveys representative sample. (Dülmer, 2007).

This is known as sampling error or accepted risk, which is fixed before the sample is taken and depends on its size. This risk is set before the sample is drawn and depends on its size: the larger the sample, the lower larger sample size, the lower the accepted risk and the better the accuracy of the survey; This technique makes it possible to use probability laws in the exploitation of data and to generalise the results and obtained on a sample to the whole population. (Lemeshow et al., 1990 ; Garthwaite et al., 2005). In fact, when interpreting the results recorded on a random sample and taking into account the accepted risk of error, the real values of the results obtained in the sample fluctuate within a certain interval (confidence interval), either around the average or around the percentage observed in this sample (Neyman & Pearson, 1928).

In medicine, the maximum accepted risk is generally 5%, and different methods are used to construct which are random sample, simple random sampling, systematic sampling, multi-stage random sampling, cluster random sampling and stratified random sampling (Smith et al., 2005). When designing a research protocol in an epidemiological study, and throughout its implementation and analysis, it is necessary to take into account factors that may lead to erroneous conclusions (Althubaiti, 2016).

For the association between a risk factor and a disease in an epidemiological study may be due to epidemiological study may be due to chance or to systematic error (bias); Bias is considered to be a particular variety of error; it is a systematic error. It is defined as a distortion in the estimate of the measure of an association between exposure to a risk factor and the occurrence of a disease and the effect may be an overestimation or underestimation of the strength of the association, or a failure to appreciate the direction of the association. (Pearce et al., 2007).

A bias results from a systematic error in the selection of subjects for a study (selection bias) or in the collection of information about the selected subjects (information bias); Biases are usually avoidable when designing a protocol and it is therefore important to anticipate the role they might play. Two broad categories of bias can be distinguished, biases that can be avoided at the time of protocol design, but also taken into account during analysis: These are known as confounding biases; And that can be avoided at the time of protocol design and study implementation, but which cannot be controlled for during analysis: these are selection bias and information bias (Tripepi et al., 2010 ; Tripepi et al., 2008).

Conclusion

Finally, it can be stated that the application and induction of mathematics, and more precisely statistics, in medicine has given a considerable step forward in research, especially in a very important field such as epidemiology.

Conflicts of Interest:

The author declares no conflicts of interest.

References

- Althubaiti, A. (2016). Information bias in health research: definition, pitfalls, and adjustment methods. *Journal of multidisciplinary healthcare*, *9*, 211.
- Anderson, S. R., Auquier, A., Hauck, W. W., Oakes, D., Vandaele, W., & Weisberg, H. I. (2009). *Statistical methods for comparative studies: techniques for bias reduction* (Vol. 170). John Wiley & Sons.
- Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the academy of marketing science*, *16*(1), 74-94.
- Bonita, R., Beaglehole, R., & Kjellström, T. (2006). *Basic epidemiology*. World Health Organization.
- Davis, F. D. (1985). *A technology acceptance model for empirically testing new end-user information systems: Theory and results* (Doctoral dissertation, Massachusetts Institute of Technology).
- Declich, S., & Carter, A. O. (1994). Public health surveillance: historical origins, methods and evaluation. *Bulletin of the World Health Organization*, *72*(2), 285.
- Dülmer, H. (2007). Experimental plans in factorial surveys: random or quota design?. *Sociological Methods & Research*, *35*(3), 382-409.
- Everitt, B., & Skrondal, A. (2002). *The Cambridge dictionary of statistics* (Vol. 106). Cambridge: Cambridge University Press.
- Garthwaite, P. H., Kadane, J. B., & O'Hagan, A. (2005). Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*, *100*(470), 680-701.
- Haux, R. (2006). Health information systems–past, present, future. *International journal of medical informatics*, *75*(3-4), 268-281.
- Hawley, D. M., & Altizer, S. M. (2011). Disease ecology meets ecological immunology: understanding the links between organismal immunity and infection dynamics in natural populations. *Functional Ecology*, *25*(1), 48-60.
- Holland, P. W. (1986). Statistics and causal inference. *Journal of the American statistical Association*, *81*(396), 945-960.
- Kim, H., Sefcik, J. S., & Bradway, C. (2017). Characteristics of qualitative descriptive studies: A systematic review. *Research in nursing & health*, *40*(1), 23-42.
- Kothari, C. R. (2004). *Research methodology: Methods and techniques*. New Age International.
- Lemeshow, S., Hosmer, D. W., Klar, J., Lwanga, S. K., & World Health Organization. (1990). *Adequacy of sample size in health studies*. Chichester: Wiley.
- Martínez-Mesa, J., González-Chica, D. A., Duquia, R. P., Bonamigo, R. R., & Bastos, J. L. (2016). Sampling: how to select participants in my research study?. *Anais brasileiros de dermatologia*, *91*(3), 326-330.
- Neyman, J., & Pearson, E. S. (1928). On the use and interpretation of certain test criteria for purposes of statistical inference: Part I. *Biometrika*, 175-240.
- Pearce, N., Checkoway, H., & Kriebel, D. (2007). Bias in occupational epidemiology studies. *Occupational and environmental medicine*, *64*(8), 562-568.
- Polgar, S., & Thomas, S. A. (2011). *Introduction to Research in the Health Sciences E-Book*.

Elsevier Health Sciences.
- Pratschke, S., Bender, A., Boesch, F., Andrassy, J., van Rosmalen, M., Samuel, U., ... & Eurotransplant Liver and Intestine Advisory Committee (ELIAC). (2019). Association between donor age and risk of graft failure after liver transplantation: an analysis of the Eurotransplant database. *Transplant International*, *32*(3), 270-279.
- Röhrig, B., Du Prel, J. B., Wachtlin, D., & Blettner, M. (2009). Types of study in medical research: part 3 of a series on evaluation of scientific publications. *Deutsches Arzteblatt International*, *106*(15), 262.
- Smith, J. L., Sturrock, H. J., Olives, C., Solomon, A. W., & Brooker, S. J. (2013). Comparing the performance of cluster random sampling and integrated threshold mapping for targeting trachoma control, using computer simulation. *PLoS Negl Trop Dis*, *7*(8), e2389.
- Tripepi, G., Jager, K. J., Dekker, F. W., & Zoccali, C. (2010). Selection bias and information bias in clinical research. *Nephron Clinical Practice*, *115*(2), c94-c99.
- Tripepi, G., Jager, K. J., Dekker, F. W., Wanner, C., & Zoccali, C. (2008). Bias in clinical research. *Kidney international*, *73*(2), 148-153.
- Trochim, W. M., & Donnelly, J. P. (2001). *Research methods knowledge base* (Vol. 2). Atomic Dog Pub..
- Wunsch, G., Russo, F., & Mouchart, M. (2010). Do we necessarily need longitudinal data to infer causal relations?. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, *106*(1), 5-18.
- Zhou, X. H., McClish, D. K., & Obuchowski, N. A. (2009). *Statistical methods in diagnostic medicine* (Vol. 569). John Wiley & Sons.