Annals of Agricultural and Environmental Medicine 2011, Vol 18, No 2, 283-284
www.aaem.pl

Internal and external validity of cohort studies

Paolo Boffetta

Mount Sinai School of Medicine, New York, USA
International Prevention Research Institute, Lyon, France

I participated with pleasure in discussions on 30 September 2011 in Warsaw during the 'Differences in Health in the Polish population' Conference, which aimed at critical assessment of first baseline results from the PONS study, a pilot prospective cohort established in Poland. The PONS study results from the 'Closing the Health Gap' Report recommending an attempt to start building in Central and Eastern Europe a network of prospective cohort studies with a biobank for the sustainable monitoring of health changes and effectiveness of health interventions in countries with much worse health indicators than Western Europe.

My assessment of the study construction is very positive. I would like to thank the colleagues from Professor Zatonski's team for their effort and the quality of the conducted research. I am also very pleased to have collaborated with my colleagues from Mount Sinai School of Medicine, NY in this study [1].

During the Conference there was a discussion on a subject that has been of interest to me for several years; therefore, here in the Speakers' Corner, I would like to present some considerations on the internal and external validity of cohort studies.

Studies of health characteristics and determinants in human populations can rarely follow the experimental paradigm, and this for various reasons. Randomization is not feasible for a number of health-related characteristics (e.g., genetic factors). In addition, it is not ethical to randomly assign for study subjects' interventions and exposures that could cause them harm. Furthermore, experimental trials, even if possible and ethical, are very complex and expensive endeavours. Health studies, more often than not, need to rely on observational approaches in which the characteristics of study subjects are measured at different points in time, and correlated among them. The most powerful of such observational investigations is the cohort study, in which a study population (a cohort) is selected and information is obtained to determine which subjects have one or more particular characteristics (e.g., carriers of a particular genetic factor), or are (or have been) exposed to an agent of interest (e.g., cigarette smoking). The study population is then followed up in time, and the incidence of disease in the exposed individuals is compared with the incidence in the unexposed. Typically, more than one exposure and more than one outcome are measured in a cohort study. The cohort study is the observational equivalent of the intervention study in that subjects are selected on the basis of their exposure status, and then followed up in time, but exposure is not allocated experimentally by the investigators.

The choice of the study population depends on the study hypotheses. In general, it is preferable to select a population in which a large proportion of subjects experience the exposure (or have the characteristics) of interest, or have a broad exposure contrast (if the exposure is common). The incidence of the main outcomes of interest also determines the choice of the study population: in general, the rarer the outcome, the larger (in terms of number of subjects and duration of the observation time) the cohort. The cohort chosen may be a general population group, such as the residents of a community, or a more narrowly defined population that can be readily identified and followed up, such as members of professional or social organizations (e.g., members of health insurance schemes, registered doctors and nurses). In general, a general population cohort allows the investigation of a broad spectrum of risk factors, but might present issues in terms of recruitment and retention during follow-up. Conversely, the advantages of a specialized population cohort are typically the high participation rate and easy implementation of the follow-up, but its members might be more homogeneous in terms of exposure, and often at lower risk for many diseases. In addition, the cohort may be selected because of high exposure to a suspected etiological factor, such as a source of ionizing radiation, a particular type of treatment (e.g., chemotherapy, radiotherapy), or an occupational hazard.

A common misconception about cohort studies is that they need to be 'representative' of some larger population from which they are drawn. This larger population is often identified as the residents of some particular geographic area, such as a city or a region. This argument is wrong for several reasons. All human populations are dynamic, and a representative sample of a larger population would soon lose its representativeness as the source population changes over time. Geographic residence is only one of many characteristics according to which a population can be defined: from a conceptual viewpoint, nothing distinguishes the ensemble of registered nurses within one country from the ensemble of people living in a given city. Furthermore, it is difficult to achieve a high participation rate and retention of a 'representative' sample of a geographically-defined population. This misconception about the need for a 'representative' cohort arises from the lack of appreciation of the difference between internal and external validity. The primary goal of a cohort study, as for any other investigation, is to generate valid results, i.e., results that do not suffer from error, except some degree of random error that can be quantified with statistical methods. The fact that the results are immediately applicable to a larger population of which the cohort is a representative sample (inhabitants of a city, registered nurses of a country) is irrelevant to the intrinsic validity of the results. This aspect of external validity (or generalizability) can be a desired characteristic of the study in order to derive immediate inference to the source population,

but it is often achieved at the cost of internal validity, because of higher opportunity for bias. Furthermore, it is possible to make inferences from a non-representative cohort to a larger population by applying weights on some basic characteristics, such as age, gender and ethnicity.

A couple of examples of classic cohort studies that are not representative of any geographically-defined population, yet have produced perfectly valid – and extremely important – results, would illustrate these theoretic arguments. Approximately 5,000 residents of the town of Framingham, in Massachusetts (USA), have been studied since 1948 [2]. The choice of the cohort was driven by logistic considerations, aimed at obtaining valid results while reducing the complexity of the study. At the time the study was set up, Framingham was a relatively stable community including both industrial and rural areas, with broad social class representation, and number of occupations and industries. The town was small enough to allow residents to come to one central examining facility and there was only one major hospital. In the study of doctors set up in England and Wales by Richard Doll and Bradford Hill in the early 1950s to assess the health effects of smoking, a postal questionnaire was sent to all doctors included in the British Medical Register who were resident in the United Kingdom [3]. The choice of this study population was determined by the facts that doctors were registered with their professional association and were therefore easy to identify and follow up, they were more likely to cooperate and the cause of death properly investigated.

REFERENCES

1. Manczuk M, Vedanthan R, Vatten L, Polewczyk A, Fuster V, Boffetta P, Zatonski W. Social Gradient of Cardiovascular Risk Factors in the sample of adults in Poland: Baseline Profile of the Polish Norwegian Study (PONS). Poster presentation at AHA Scientific Sessions 2011, Florida, USA, 12-14 Nov 2011.
2. Dawber TR, Meadors GE, Moore EE. Epidemiological approaches to heart disease: the Framingham Study. Am J Public Health 1951;41:279-86.
3. Doll R, Hill AB. The mortality of doctors in relation to their smoking habits. A preliminary report. Br Med J 1954;i:1451-5.