FACTA UNIVERSITATIS
Series: **Working and Living Environmental Protection** Vol. 13, N° 2, 2016, pp. 113 - 119

CASE-CONTROL DESIGN AS A METHOD IN ASSESSING ENVIRONMENTAL EXPOSURES AND HEALTH OUTCOMES

UDC 614.7:502

Biljana Kocić, Aleksandra Stanković, Mirjana Aranđelović

University of Niš, Faculty of Medicine, Niš, Serbia

Abstract. *Case-control study has made, and will continue to make, important contributions to the environmental health field. Case-control studies are very common in environmental epidemiologic research as hypothesis-testing designs. The purpose of this article is to discuss the principles of a case-control study design and related methodological issues in environmental epidemiology. Case-control studies assessing environmental risk factors are pursued due to low cost, efficiency, disease latency, etc. Weaknesses do not negate strengths of the case-control design. Many examples illustrate the value of the case-control study as investigative approach in assessing environmental exposures and health outcomes. Recognizing that well-designed case-control studies can provide valid results is important, so that investigators can both critically appraise and appropriately design case-control studies to address important environmental research questions.*

Key words: *case-control study, environmental exposures, health outcomes*

1. INTRODUCTION

Epidemiology is one of the core disciplines used to examine the associations between environmental hazards and health outcomes. Study designs can be arranged on a continuum ranging from hypothesis-generating designs that provide limited information to complex hypothesis-testing designs. For the particular problem being investigated, some designs are better than others, depending upon what is to be achieved, the availability of study populations, the disease or health outcome studied, and the need to uncover disease etiology. Case-control studies are usually the most readily and cheaply carried out of all analytic epidemiologic studies. They are very common in environmental epidemiologic research.

The purpose of this article is to discuss the principles of a case-control study design and related methodological issues in environmental epidemiology. The focus is on examples aimed at evaluating causal hypotheses regarding exposures to suspected health hazards.

Received October 16, 2014 / Accepted May 27, 2016
Corresponding author: Biljana N. Kocić
University of Niš, Faculty of Medicine, Dr. Zoran Djincica 81, 18000 Niš, Serbia
E-mail: biljaizzz@yahoo.com

2. Study Design and Methodological Issues

Case-control studies are observational because no intervention is attempted and no attempt is made to alter the course of the disease. Case- control study compares patients who have a disease or outcome of interest (cases) with patients who do not have the disease or outcome (controls), and looks back retrospectively to compare how frequently the exposure to a risk factor is present in each group to determine the relationship between the risk factor and the disease. Cases may be patients from hospitals, specialized clinics, medical practices, or disease registries such as cancer registries. Controls can be from the community, healthy persons, friends or relatives of the cases, or patients from hospitals or clinics, affected by a disease that is etiologically unrelated to the outcome of interest. Cases and controls are generally matched according to criteria such as sex, age, race or other criteria, because they are typically strong confounders of disease [17]. Confounders are variables associated with the risk factor and may potentially be a cause of the outcome [5].

Case-control studies are designed to estimate odds ratio (OR), as a measure of association between the exposure and outcome [8]. The exposure- OR refers to "the ratio of odds in favor of exposure among cases to the odds in favor of exposure among the controls [11]. An odds ratio, statistically significant, of more than 1 suggests a positive association between the exposure and disease or other outcome. One method to determine past exposure is to interview subjects from both groups regarding their exposure history. It may be possible to conduct direct measurements of the environment for various types of exposures.

Case-control studies are well suited to investigate rare outcomes or outcomes with a long latency period because subjects are selected from the outset by their outcome status. Thus, case-control studies are quick, relatively inexpensive to implement, require comparatively fewer subjects, and allow for multiple exposures or risk factors to be assessed for one outcome [9,4]. Limitations of any case-control study include small sample size, selection bias, environment exposure misclassification, recall bias, and confounding. This suggests that a blended approach that combines questionnaires, exposure modeling (e.g., using residence information to evaluate past air pollution exposure), biological measurements (e.g., bone or blood lead measurements), and possibly environmental monitoring data (e.g., water quality measurements) could provide greater specificity of exposures and accuracy. Recall bias is also a limitation when using questionnaires, especially when seeking long term or historical information regarding exposure as these and other factors of possible significance may have occurred many years before the diagnosis and study enrollment, and thus participants may have difficulty remembering potential data. These biases decrease the internal validity of the investigation and should be carefully addressed and reduced in the study design. Investigations examining rare outcomes may have a limited number of cases to select from, whereas the source population from which controls can be selected is much larger. In such scenarios, the study may be able to provide more information if multiple controls per case are selected. This method increases the "statistical power" of the investigation by increasing the sample size. The precision of the findings may improve by having up to about three or four controls per case [16,18].

3. Examples of the Environmental Epidemiology Literature

As we have noted, selecting the appropriate group of controls can be one of the most demanding aspects of a case-control study. An important principle is that the distribution of exposure should be the same among cases and controls; in other words, both cases and controls should stem from the same source population. The investigator may also consider the control group to be an at-risk population, with the potential to develop the outcome. Because the validity of the study depends upon the comparability of these two groups, cases and controls should otherwise meet the same inclusion criteria in the study. A case-control study design that exemplifies this methodological feature is by Chung and colleagues, who examined maternal cigarette smoking during pregnancy and the risk of newborns developing cleft lip/palate [1]. A salient feature of this study is the use of the 1996 U.S. Natality database, a population database, from which both cases and controls were selected. This database provides a large sample size to assess newborn development of cleft lip/palate (outcome), which has a reported incidence of 1 in 1000 live births [2], and also enabled the investigators to choose controls (i.e., healthy newborns) that were generalizable to the general population to strengthen the study's external validity. A significant relationship with maternal cigarette smoking and cleft lip/palate in the newborn was reported in this study (adjusted OR 1.34, 95% CI 1.36-1.76) [1].

An example of a case-control investigation is by Zhang and colleagues who examined the association of environmental and genetic factors associated with rare congenital microtia [20], which has an estimated prevalence of 0.83 to 17.4 in 10,000 [13]. They selected 121 congenital microtia cases based on clinical phenotype, and 152 unaffected controls, matched by age and sex in the same hospital and same period. Controls were of Hans Chinese origin from Jiangsu, China, the same area from where the cases were selected. This allowed both the controls and cases to have the same genetic background, important to note given the investigated association between genetic factors and congenital microtia. To examine environmental factors, a questionnaire was administered to the mothers of both cases and controls. The authors concluded that adverse maternal health was among the main risk factors for congenital microtia, specifically maternal disease during pregnancy (OR 5.89, 95% CI 2.36-14.72), maternal toxicity exposure during pregnancy (OR 4.76, 95% CI 1.66-13.68), and resident area, such as living near industries associated with air pollution (OR 7.00, 95% CI 2.09-23.47) [20]. A case-control study design is most efficient for this investigation, given the rarity of the disease outcome. Because congenital microtia is thought to have multifactorial causes, an additional advantage of the case-control study design in this example is the ability to examine multiple exposures and risk factors.

A case-control study of Teo et al. (2006) [14], with 27,089 participants (12,461 cases, 14,637 controls), in 52 countries looked at the relation between risk of acute myocardial infarction (AMI) and current or former smoking, type of tobacco, amount smoked, effect of smokeless tobacco, and exposure to secondhand smoke (SHS). Current smoking was associated with a greater risk of non-fatal AMI odds ratio (OR=2.95, 95% CI 2.77-3.14, $p<0.0001$) compared with never smoking; risk increased by 5.6% for every additional cigarette smoked. The OR associated with former smoking fell to 1.87 (95% CI 1.55-2.24) within 3 years of quitting. A residual excess risk remained 20 or more years after quitting (1.22, 1.09-1.37). Exclusion of individuals exposed to SHS in the never smoker reference group raised the risk in former smokers by about 10%. Smoking beedies alone (indigenous to South Asia) was associated with increased risk (2.89, 2.11-3.96) similar to that associated with cigarette smoking. Chewing tobacco alone was associated with OR 2.23 (1.41-3.52), and smokers who also chewed tobacco had the highest increase in risk (4.09, 2.98-5.61). SHS was associated with a graded increase in risk related to exposure; OR was 1.24 (1.17-1.32) in individuals who were least exposed (1-7 h per week) and 1.62 (1.45-1.81) in people who were most exposed (>21 h per week). Young male current smokers had the highest population attributable risk (58.3%; 95% CI 55.0-61.6) and older women the lowest (6.2%, 4.1-9.2). Population attributable risk for exposure to SHS for more than 1 h per week in never smokers was 15.4% (12.1-19.3).

Another important example of environmental research, especially regarding exposure data obtaining issues, is a case-control study among female residents of Long Island, New York, examined the possible association between exposure to electromagnetic fields (EMFs) and breast cancer [12]. Subjects were those who were younger than 75 years of age, who lived in the study area for 15 years or longer, and were identified between August 1, 1996, and June 20, 1997. Cases (n=576) consisted of women diagnosed with in situ or invasive breast cancer. Controls (n=585) were selected from the same community by random digit dialing procedures. In-home data collection included various spot and 24-hour EMF measurements, ground-current magnetic field measurements, wire mapping of overhead power lines servicing the home, and an interview. Odds ratios and 95% confidence intervals were based on multivariate logistic regression analyses. All odds ratios were close to 1 and non-significant. For the highest quartile of 24-hour EMF measurements, the odds ratio was 0.97 (95% confidence interval (CI): 0.69, 1.37) in the bedroom and 1.09 (95% CI: 0.78, 1.51) in the most lived-in room. For the highest exposure category of ground-current measurements, the odds ratio was 1.13 (95% CI: 0.88, 1.44) in the bedroom and 1.08 (95% CI: 0.85, 1.38) in the most lived-in room. These and other EBCLIS results agree with other recent reports of no association between breast cancer and residential EMF exposures.

Case-control studies have played a vital role in the development of many fruitful lines of study. For example, the relationship of cigarette smoking to lung cancer was demonstrated in case-control studies before any cohort studies of this question were carried out [3]. Because of their relatively low cost, case-control studies should often be the first approach to testing of a hypothesis. They provide an excellent way to investigate whether any of several exposures is associated with a particular disease. This feature may facilitate an exploratory study (sometimes referred to as a “fishing expedition”) to find clues and leads for further study [7]. For example, Doll and Hill's study in 1950 [3], has come to be viewed as a model of the case-control investigation. Notification of cancer cases (lung, colon, stomach, rectum) were received from 20 London hospitals. Each case was interviewed by a social worker who was also “instructed to interview a patient of the same sex, within the same five-year age group, and in the same hospital at or about the same time” who did not have cancer. Attention was paid to the duration of smoking, to histories of starting and stopping smoking, and to the amount smoked. This study devised the convention of setting the lower threshold for lifetime smoking at one cigarette per day for a year. A six-month re-interview of a subset of subjects showed remarkable consistency in self-reported smoking histories. Contrast were made between cases of lung cancer and matched controls in overall smoking, in amount smoked most recently, in maximum ever smoked, in age of onset of smoking and in duration of smoking. Pipe smoking was shown to have a weaker relationship to lung cancer than cigarette smoking. Stratified analyses were used to deal with potential confounders, including urban/rural residence, cancer diagnosis of controls and potential interviewer bias. Unlike any other case-control study of the period, Doll and Hill used the distribution of smoking in lung cancer patients to develop "ratios" for lung cancer risk in London smokers, assuming a smoking distribution that paralleled that of the control population. This yielded estimates of relative risk for lung cancer from smoking 10, 20 and 60 cigarettes per day of 19, 26 and 65; odds ratios were not calculated.

Another cost-saving application is the nested case-control study. Nested case-control studies within retrospective cohorts have come to be an important tool of epidemiologists, as in Thériault et al.'s study of the occupational hazards of leukemia and brain cancer associated with exposure to 60 Hz electromagnetic fields produced by the passage of electricity [15]. Investigators carried out a cohort study of 223 000 men employed by Electricité de France-Gaz de France, Ontario Hydro, and Hydro-Québec, who were observed during the period 1970 to 1989. Cases of cancer were ascertained in France from company medical records during employment. In Canada they were ascertained by matching the names of employees with the records of local cancer registries and, in Quebec, from company medical records and death certificates. Over 4000 men were identified as having developed cancer. To compare with them, controls were selected randomly from members of the whole population matched for the same utility, year of birth, and alive on the date the affected employee developed cancer: four for each man with one of the cancers of special interest (defined in this study as any haematopoetic cancer, brain cancer, or melanoma) and one for each man with a cancer of another type. Full occupational histories were built up for all members of both groups from company records. Measurements of the fields to which a sample of over 2 000 men currently employed were exposed were then made by means of personal dosimeters worn throughout a five-day week and jobs with similar mean exposures were collapsed into between 32 and 65 occupational groups, the number of groups varying between the different utilities. Allowance was made for procedural and power changes over the period of the men's exposures and each man's exposure was expressed in "μ T years" by multiplying the mean exposure in each occupational group by the time spent in it. Odds ratios, adjusted for socio-economic status, were then calculated for 31 cancer types for men with exposures equal to or above the median and exposures equal to or above the 90th percentile in comparison with men whose exposures were less than the median. Only three of the cancers showed statistically significant excesses in one or other of the more heavily exposed groups: those for acute non-lymphoid leukemia, acute myeloid leukemia, and astrocytoma. Though not conclusive, these findings greatly strengthen belief in the idea that 60 Hz electromagnetic fields may cause occupational hazards of these two diseases. It was a major task to compute the occupational exposures of the 11000 men selected for a nested case-control study and it would have been quite impracticable, at the same level of detail, to have done so for 20 times as many men, which a simple cohort study would have required.

In Michigan, a case-control study was conducted to explore the role of environmental factors in the development of amyotrophic lateral sclerosis (ALS) [19]. Sixty-six cases and 66 age- and gender-matched controls were recruited. Detailed information regarding residence history, occupational history, smoking, physical activity, and other factors was obtained using questionnaires. The association of ALS with potential risk factors, including smoking, physical activity and chemical exposure, was investigated using conditional logistic regression models. As compared to controls, a greater number of our randomly selected ALS patients reported exposure to fertilizers to treat private yards and gardens and occupational exposure to pesticides in the last 30 years than our randomly selected control cases. Smoking, occupational exposures to metals, dust/fibers/fumes/gas and radiation, and physical activity were not associated with ALS when comparing the randomly selected ALS patients to the control subjects. To further explore and confirm results, exposures over several time frames, including 0–10 and 10–30 years earlier, were considered, and analyses were stratified by age and gender. Pesticide and fertilizer exposure were both significantly associated with ALS in the randomly selected ALS patients. While study results need to be interpreted cautiously given the small sample size and the lack of direct exposure measures, these results suggest that environmental and particularly residential exposure factors warrant close attention in studies examining risk factors of ALS. Case-control studies assessing environmental risk factors in ALS are pursued due to low cost, efficiency, disease latency, and tendency to affect older individuals [6,10].

4. Conclusion

Case-control studies are usually the most readily and cheaply carried out of all analytic epidemiologic studies. Many examples illustrate the value of the case-control study. For rare diseases they may be the only practical approach. It is a relatively rapid and reliable method of establishing evidence of an association between an exposure to a risk (or protective) factor and an unfavorable (or favorable) outcome. It can study several possible causes or exposures to risk simultaneously. It does not require study of large numbers-usually requires only a few cases. It is an excellent way to study diseases with long latency. It can often make use of existing records. It is obvious that case-control studies have played a vital role in the development of many fruitful lines of study.

Case-control study has made, and will continue to make, important contributions to the environmental health field. Case-control studies assessing environmental risk factors are pursued due to low cost, efficiency, disease latency, etc. Weaknesses do not negate strengths of the case-control design. Many examples illustrate the value of the case-control study as investigative approach in assessing environmental exposures and health outcomes. Recognizing that well-designed case-control studies can provide valid results is important, so that investigators can both critically appraise and appropriately design case-control studies to address important environmental research questions.

Acknowledgement: *This work was supported by grant No III 43014, from the Ministry of Education, Science and Technological Development of the Republic of Serbia.*

References

1. Chung, K.C., Kowalski, C.P., Kim, H.M. and Buchman, S.R.: Maternal cigarette smoking during pregnancy and the risk of having a child with cleft lip/palate. Plast. Reconstr. Surg., Vol. 105, pp. 485–491, 2000.
2. Das, S.K., Runnels, R.S. Jr, Smith, J.C. and Cohly H.H.: Epidemiology of cleft lip and cleft palate in Mississippi.South. Med. J., Vol.88, pp. 437–442, 1995.
3. Doll, R. and Hill, A.B.: Smoking and carcinoma of the lung: preliminary report. Br Med J, ii:pp. 739-748, 1950.
4. Elwood, M.: Critical Appraisal of Epidemiological Studies and Clinical Trials. 3rd Ed. Oxford University Press; Oxford: pp. 1-570, 2007.
5. Everitt B.S. and Palmer C.R.: Encyclopaedic Companion to Medical Statistics. Hodder Arnold; London:2005.

6. Fang F, Kwee LC, Allen KD, Umbach DM, Ye W, Watson, M., Keller, J., Eugene Z. Oddone, E. Z., Sandler, D. P., Schmidt S. and Kamel, F.: Association between blood lead and the risk of amyotrophic lateral sclerosis. Am. J. Epidemiol., Vol.171: 1126–1133, 2010.
7. Friedman, G.D.: Primer of epidemiology, 5th ed. McGraw-Hill Companies, NY, USA: pp. 128-129, 2004.
8. Greenberg R.S. and others (eds).: Medical epidemiology, (LANGE Basic Science) 4th ed. McGraw-Hill Medical, pp. 147-161, 2004.
9. Hulley, S.B. and others: Designing Clinical Research: An Epidemiologic Approach. 2nd Ed. Lippincott Williams & Wilkins; Philadelphia: pp. 1-336, 2001.
10. Noonan, C.W., Reif, J.S., Yost, M. and Touchstone, J. Occupational exposure to magnetic fields in case-referent studies of neurodegenerative diseases. Scand. J. Work. Environ. Health., 28, pp. 42–48, 2002.
11. Porta M.A.: Dictionary of Epidemiology. 5th ed. New York, NY: Oxford University Press; 2008.
12. Schoenfeld, E.R., O'Leary, E.S., Henderson, K., Grimson, R., Kabat, G.C., Ahnn, S., Kaune, W.T., Gammon, M.D. and Leske, M.C.: Electromagnetic fields and breast cancer on Long Island: a case-control study. Am. J. Epidemiol., Vol. 158, pp. 47-58, 2003.
13. Suutarla, S., Rautio, J., Ritvanen, A., Ala-Mello, S., Jero, J. and Klockars T.: Microtia in Finland: Comparison of characteristics in different populations. Int. J. Pediatr. Otorhinolaryngol., Vol. 71, pp. 1211–1217, 2007.
14. Teo, K. K., Ounpuu, S., Hawken, S., Pandey, M., Valentin, V., Hunt, D., Diaz, R., Rashed, W., Freeman, R., Jiang, L., Zhang, X. and Yusuf, S.: Tobacco use and risk of myocardial infarction in 52 countries in the INTERHEART study: A case-control study. Lancet, Vol. 368, pp. 647-658, 2006.
15. Thériault, G., Goldberg, M., Miller, A.B, Armstrong, B., Guénel, P., Deadman, J., Imbernon, E., To, T., Chevalier, A., Cyr, D. and others: Cancer risks associated with occupational exposure to magnetic fields among electric u.tility workers in Ontario and Quebec, Canada, and France, 1970–1989. Am. J. Epidemiol., Vol. 139, pp. 550–572, 1994.
16. Ury, H.K.: Efficiency of case-control studies with multiple controls per case: Continuous or dichotomous data. Biometrics, Vol.31, pp. 643–649, 1975.
17. Wacholder S., McLaughlin J.K., Silverman D.T. and Mandel J.S.: Selection of controls in case-control studies. I. principles. Am. J. Epidemiol., 135:pp.1019–1028, 1992.
18. Woodward, M.: Epidemiology: Study Design and Data Analysis. 2nd Ed. Chapman & Hall/CRC; London: pp. 1-849, 2005.
19. Yu, Y., Su, F-C., Callaghan, B.C., Goutman, S.A., Batterman, S.A. and Feldman, E.L.: Environmental Risk Factors and Amyotrophic Lateral Sclerosis (ALS): A Case-Control Study of ALS in Michigan. 2014. PLoS ONE 9(6): e101186. doi:10.1371/journal.pone.0101186
20. Zhang Q.G., Zhang, J., Yu, P. and Shen, H.: Environmental and genetic factors associated with congenital microtia: A case-control study in Jiangsu, China, 2004 to 2007. Plast. Reconstr. Surg., Vol. 124, pp.1157–1164, 2009.

ANAMNESTČKA STUDIJA KAO METOD ZA SAGLEDAVANJE IZLOŽENOSTI U ŽIVOTNOJ SREDINI I ZDRAVSTVENIH ISHODA

Anamnestičke studije su doprinele, i nastaviće da daju značajan doprinos istraživanju u oblasti zaštite zdravlja i izloženosti noksama u životnoj sredini. Ove studije se često koriste, pre svega za testiranje etioloških hipoteza. Cij rada je bio da sagleda karakteristike dizajna anamnestičke studije, i razmotri metodološke principe na primerima izloženosti u životnoj i radnoj sredini i zdravstvenih ishoda. Uprkos nedostacima, mnogobrojni primeri iz literature ukazuju na značaj ovog istraživačkog dizajna, pre svega zbog efikasnosti, mogućnosti da istražuje uticaj brojnih faktora na određeni poremećaj zdravlja, zbog angažovanja relativno malog broja ispitanika i sredstava, mogućnosti da istraži poremećaji zdravlja sa niskom učestalošću kao i onih s dugim latentnim periodom.

Ključne reči: *anamnestička studija, izloženosti u životnoj sredini, zdravstveni ishodi*