UNDERGRADUATE RESEARCH IN NATURAL AND CLINICAL SCIENCE AND TECHNOLOGY (URNCST) JOURNAL
Read more URNCST Journal articles and submit your own today at: https://www.urncst.com

RESEARCH METHODS PRIMER OPEN ACCESS

Longitudinal Prospective Cohort Studies: A Research Method Primer

Jasmine Zouhair, MD Student [1]*

[1] Department of Medicine, International University of the Health Sciences, Winnipeg, Manitoba, Canada R3P 0N5

*Corresponding Author: jasminez@iuhs.edu

Abstract

Introduction: A cohort study is a nonexperimental study design used to investigate the outcomes of a particular risk factor. They help researchers understand the prevalence, distribution, and correlation of variables in a population and function by following participants over a period of time, typically years.

Utility: The results of well-designed observational studies are comparable to those of randomized controlled trials. Among some of the strengths of cohort studies is the ability to measure incidence rates, and to allow for a wide range of variables to be examined. In addition, one can examine disease progression and natural history because of their longitudinal design characteristic. Specifically, they are advantageous for rare exposures since subjects are chosen based on exposure status and can be monitored throughout the study for any changes caused by said exposure. Although they can infer a relationship between variables, they do not confirm causality.

Challenges: One of the greatest challenges posed by cohort studies is the considerable amount of time and funding required to conduct them as they require large samples. Other challenges include, but are not limited to, maintaining follow-ups and accounting for withdrawals, and minimal control over the variables that are being studied

Limitations: Variables may be measured incorrectly or inconsistently, resulting in information bias. For diseases with extensive latency periods, this study strategy is ineffective and cannot be used to establish causation between variables because the disease may have not completely manifested in the time it takes to conduct the study. Another significant limitation of this design is the sources of bias that could jeopardize the reliability of the study as a result of faulty measurement, an unrepresentative sample, or the differing impact of other factors on the association of interest.

Keywords: Cohort study; prospective; longitudinal cohort study; risk factor; exposure; outcomes

Introduction

Cohort studies are nonexperimental studies in which groups of individuals are studied over a prolonged period of time, often years, based on the common characteristics they share, as well as exposure to risk factors relevant to the subject of study [1]. In prospective cohort studies, data is collected prospectively with the investigator following individuals possessing the variable of interest. The investigator can gather data on additional factors that are relevant to the research question, which enables them to further explore the impact of the variable and measure the outcomes [1,2]. However, it is not always the case that a risk factor exists in a cohort, since there are many healthy cohorts that can be studied. Epidemiological and clinical studies have taken advantage of this method in a variety of ways such that, in epidemiology, they aid in understanding what variables may enhance or decrease the likelihood of contracting a disease [2]. In clinical research, cohort studies are used when there is evidence that implies a clear correlation between an exposure or risk factor and an outcome [3]. Importantly, we can use them to determine the causes, prognosis, and incidences of certain variables. Due to their longitudinal nature, a follow-up stage is initiated in which outcomes are measured to assess the importance of specific risk factors and their impact [2]. Despite providing valuable insights, loss of follow-up is one of the major disadvantages of cohort studies. Although inherent, this loss often leads to biases in the final results and may affect the statistical significance of said results [5]. Strategies, such as regular schedules of follow-ups throughout various stages of the study as opposed to one follow-up at the end of the study have been undertaken to mitigate large losses of follow-ups [6]. Cost and time are other disadvantages, however prospective cohort designs are still effective for obtaining valuable information in epidemiological and clinical studies.

Utility

Clinical Utility

One of the greatest strengths of cohort studies is the information they provide on how exposure can be associated with certain outcomes, such as in the case of a disease. They provide a greater degree of information pertaining to

potential causal relationships due to their temporal design [7]. Another major strength is their accuracy in being applied to a broader population outside of the study as well-done cohort studies may provide results similar to those of experimental studies [4]. Moreover, investigators have the ability to study multiple exposures and outcomes in one study [8]. As such, they have been utilized in different academic disciplines. For example, the well-known Framingham Heart Study, which began in 1948, amassed 5,209 men and women between the ages of 30-62 from Framingham, Massachusetts [9]. Investigators sought to determine if there were any associations between lifestyle and cardiovascular disease development. Every two to six years, participants returned to provide investigators with a medical report and undergo physical exams. There have been two more cohorts since then; one in 1971, the second generation of the first cohort and another in 2002, the third generation of the first cohort. Their finding was that certain lifestyles are correlated with an increased risk for heart disease. Investigators were able to use their data to find plausible associations between cardiovascular disease development and lifestyle choices, though it is important to keep in mind that longitudinal studies are not conclusive. This demonstrates the way in which cohort studies can provide a wealth of data in particular fields of study despite no experimental procedures being conducted. Compared to experimental study designs, cohort studies are observational in nature, whereby investigators do not interfere but rather observe and evaluate [4].

The Nurses' Health study is another example of a prospective cohort study, comprised of 121,700 participants that began in 1980 and is ongoing [10]. Once again, investigators gathered data from its participants in order to examine the relationship between reproductive, dietary, hormonal, and lifestyle factors and coronary heart disease in female registered nurses. A more recent birth cohort study is the Millennium Cohort Study (MCS) which has been following approximately 19,000 people born in the years 2000-2002 from across the UK [11]. Their goal is to assess the cohort members' physical, behavioural, and socio-economical development over time. These three studies have in common their capacity to identify outcomes based on particular risk factors. Moreover, the fact that they are observational has allowed investigators to draw strong associations between variables of interest [2]. Specifically, the collection of data over regular intervals reduces recall error, which is why these large-scale studies have been so successful.

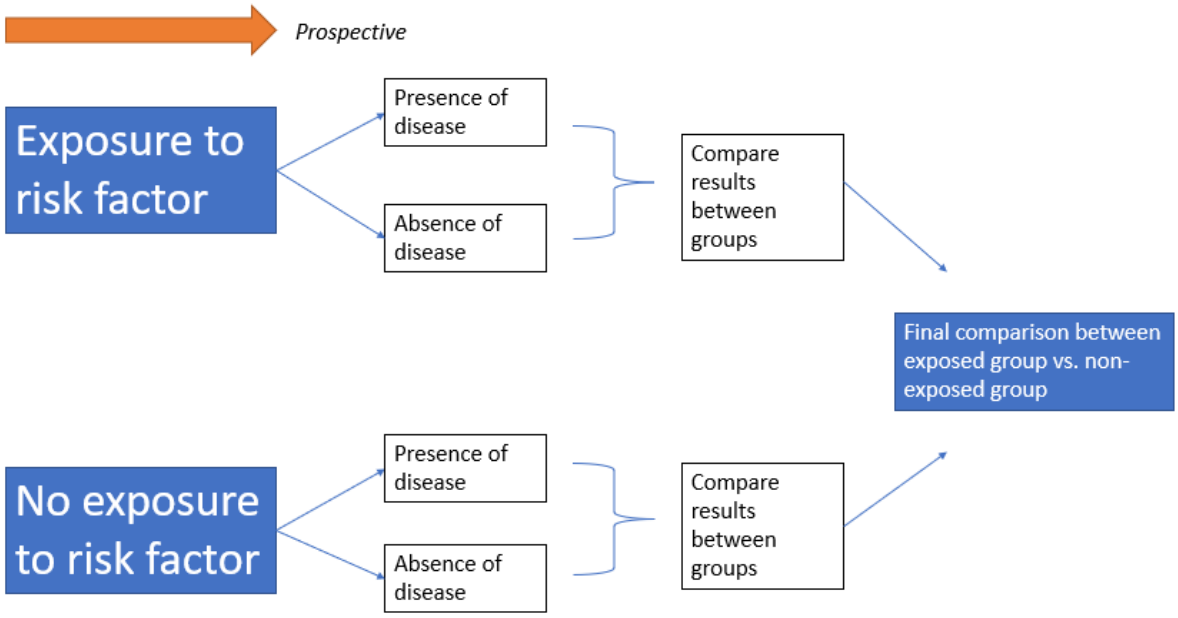

Figure 1. Prospective cohort study design (created with Microsoft PowerPoint).

Challenges

Despite the well-defined utilities of cohort studies, challenges do arise as an inherent part of the study design. Prospective cohort studies can be a better option than randomized control trials in cases where it is unethical to expose participants to risk factors that could possibly affect their health [12]. However, challenges such as the considerable amount of money and time required to conduct them can be overlooked. For instance, the National Children's Study (NCS) in the USA, a birth cohort study which has recruited at least 100 000 patients and has a follow-up period of 21 years estimated its costs to be 2.7 billion USD [13]. In Japan, Mishiro and colleagues [14] reported that the cost of a 4-year baseline survey was 153

million yen or approximately 1.4 million CAD. In their calculations, they included costs for advertisement, patient recruitment, material fees, and equipment. Although the total cost may vary depending on the length and number of participants in the study, in the model survey they conducted in order to provide a close estimate of the required expenses, they calculated their total cost to be 14.4 million yen per year (136 814.40 CAD), and this only took into account the minimum number of items required to enrol 7400 participants.

Following participants for a long period of time is another key challenge presented by cohort studies. This can lead to loss of patient follow up and affect the outcome of the results, and create biases that could possibly negatively affect their health [5]. Specifically, self-selection bias, arising in cases where other criteria except random sampling is used to sample a population of interest [15], can skew the estimates of exposure-outcome relationships [16]. This is due to the fact that participants may select themselves into a group and thus, give incorrect information during the study [17].

Loss of follow up reduces the power of statistical analyses and in doing so, results in an underestimation or overestimation of outcomes because of incomplete data [18]. That is why it is recommended that contact information for the patient be recorded as soon as the patient is entered into the study [19]. Furthermore, it could be the case that participants alter their behaviour during the study because they are aware that they are being observed [2]. Data analysis, too, can prove complex due to the vast amount of information being collected per patient throughout the trial and can distort results [2].

Limitations

Prospective studies cannot ensure that the study groups are comparable at the start of the investigation [20], and thus, variables may be measured incorrectly or inconsistently. Information bias is one of the most frequent kind of biases that undermines the reliability of health studies and derives from the method used to acquire or validate research measurements [21]. However, several approaches can be taken to mitigate informational bias, such as independent checks on researchers who hold differing opinions or the use of blinding or masking procedures [21] to withhold information that may influence study results. Additionally, other confounding variables may develop during the study that the investigators might not be completely aware of. This renders cohort studies less effective for establishing causation between variables. Thus, we can only credibly attribute an observed effect to be causative in randomised trials and other investigations because of incomplete and potentially biased data [20].

Other limitations include the selection of an unrepresentative population or the differing impact of other factors on the association of interest [20]. An example of this is the healthy-worker effect, a common type of selection bias wherein researchers fail to choose a proper comparison group composed of healthy and unhealthy people [22]. Since children, sick people, and elderly retired people cannot get jobs in the workforce and are considered 'unhealthy', it implies that people in the workforce are 'healthier'. This bias highlights how critical it is for investigators to select representative populations and be aware of confounding factors when measuring degrees of association. To ensure that investigators are mitigating bias and confounding factors, several important questions to keep in mind are: what methods were used to assess and control for confounding? And did investigators gather information on potential confounding factors? [23].

Sample size is yet another criterion to take into account. Sample size must not be too large or too little, and because results have to be extrapolated from these studies and applied to the general population, it is important that all the information gathered is relevant to the question being investigated [24]. Sample sizes that are too large can give rise to sampling bias, ascertainment bias, and measurement errors, to list a few [25]. Very small sample sizes may alter the generalizability of the results, which is a clear limitation. Moreover, it can lead to repetitive measurements and skewed normal distributions [26].

Conclusions

Cohort studies can be a nonexperimental, alternative method useful for gathering data and determining multiple outcomes for a given exposure. Due to their longitudinal design, they enable researchers to observe exposures over a period of time to monitor any changes that may occur to patients, as well as collect data that can be used to infer a relationship between any exposure and a possible outcome. However, cohort studies are costly, prone to bias, and prone to loss of follow-up. Investigators may take steps to limit bias and loss of follow-up, but other disadvantages exist, such as sample sizes and their application to long latency diseases. Nevertheless, cohort studies have the potential to provide accurate results of outcomes, attesting to their traction in clinical and epidemiological studies.

List of Abbreviations Used

HEALS: Health Effects of Arsenic Longitudinal Study
MCS: Millennium Cohort Study
NCS: National Children's Study

Conflicts of Interest

The author(s) declare that they have no conflict of interests.

Ethics Approval and/or Participant Consent

This paper is a research methods primer and did not require ethical approval or participant consent.

Authors' Contributions

JZ: made substantial contributions to review planning, assisted with the collection and organization of primary data, and drafted the manuscript.

Acknowledgements
We would like to thank Pallavi Dutta for providing draft revision.

Funding
This study was not funded.

References

[1] Setia MS. Methodology series module 1: cohort studies. Indian journal of dermatology. 2016 Jan;61(1):21. https://doi.org/10.4103/0019-5154.174011.
[2] Barrett D, Noble H. What are cohort studies?. Evidence-based nursing. 2019 Oct 1;22(4):95-6. https://doi.org/10.1136/ebnurs-2019-103183.
[3] Wang X, Kattan MW. Cohort studies: design, analysis, and reporting. Chest. 2020 Jul 1;158(1):S72-8. https://doi.org/10.1016/j.chest.2020.03.014
[4] Song JW, Chung KC. Observational studies: cohort and case-control studies. Plastic and reconstructive surgery. 2010 Dec;126(6):2234. https://doi.org/10.1097/PRS.0b013e3181f44abc.
[5] Kristman V, Manno M, Côté P. Loss to follow-up in cohort studies: how much is too much?. European journal of epidemiology. 2004 Aug;19(8):751-60. https://doi.org/10.1023/b:ejep.0000036568.02655.f8.
[6] Manolio TA, Bailey-Wilson JE, Collins FS. Genes, environment and the value of prospective cohort studies. Nature Reviews Genetics. 2006 Oct;7(10):812-20. https://doi.org/10.1038/nrg1919.
[7] Gilmartin-Thomas JF, Liew D, Hopper I. Observational studies and their utility for practice. Aust Prescr. 2018 Jun;41(3):82-85. https://doi.org/10.18773/austprescr.2018.017.
[8] Euser AM, Zoccali C, Jager KJ, Dekker FW. Cohort studies: prospective versus retrospective. Nephron Clinical Practice. 2009;113(3):c214-7. https://doi.org/10.1159/000235241.
[9] Framingham heart study (FHS) [Internet]. Nih.gov. [cited 2022 Jun 2]. Available from: https://www.nhlbi.nih.gov/science/framingham-heart-study-fhs.
[10] Nurses' Health Study (Cardiovascular Component) [Internet]. Clinicaltrials.gov. [cited 2022 Jun 2]. Available from: https://www.clinicaltrials.gov/ct2/show/NCT00005152.
[11] Smith K, Joshi H. The millennium cohort study. Popul Trends. 2002 Spring;(107):30-4.
[12] Ahsan H, Chen Y, Parvez F, Argos M, Hussain AI, Momotaj H, Levy D, Van Geen A, Howe G, Graziano J. Health effects of arsenic longitudinal study (HEALS): description of a multidisciplinary epidemiologic investigation. Journal of exposure science & environmental epidemiology. 2006 Mar;16(2):191-205. https://doi.org/10.1038/sj.jea.7500449.
[13] Doyle A, Golding J. The costing and funding of longitudinal birth cohort studies. Paediatric and Perinatal Epidemiology. 2009 Jul;23:86-92. https://doi.org/10.1111/j.1365-3016.2008.01011.x.
[14] Mishiro I, Sawada N, Iwasaki M, Ohashi K, Tsugane S. Minimum cost estimation of a baseline survey for a molecular epidemiology cohort study: collecting participants in a model region in Japan. Journal of epidemiology. 2016 Oct 5;26(10):522-9. https://doi.org/10.2188/jea.je20150163.
[15] Heckman JJ. Selection bias and self-selection. InMicroeconometrics 2010 (pp. 242-266). Palgrave Macmillan, London. https://doi.org/10.1057/9780230280816_29.
[16] Biele G, Gustavson K, Czajkowski NO, Nilsen RM, Reichborn-Kjennerud T, Magnus PM, Stoltenberg C, Aase H. Bias from self selection and loss to follow-up in prospective cohort studies. European journal of epidemiology. 2019 Oct;34(10):927-38. https://doi.org/10.1007/s10654-019-00550-1.
[17] Elston DM. Participation bias, self-selection bias, and response bias. Journal of the American Academy of Dermatology. 2021 Jun 18. https://doi.org/10.1016/j.jaad.2021.06.025.
[18] Howe LD, Tilling K, Galobardes B, Lawlor DA. Loss to follow-up in cohort studies: bias in estimates of socioeconomic inequalities. Epidemiology. 2013 Jan;24(1):1-9. https://doi.org/10.1097/EDE.0b013e31827623b1.
[19] Mann CJ. Observational research methods. Research design II: cohort, cross sectional, and case-control studies. Emergency medicine journal. 2003 Jan 1;20(1):54-60. https://doi.org/10.1136/emj.20.1.54.
[20] Healy P, Devane D. Methodological considerations in cohort study designs. Nurse Res. 2011;18(3):32-6. https://doi.org /10.7748/nr2011.04.18.3.32.c8461.
[21] Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. Journal of multidisciplinary healthcare. 2016;9:211. https://doi.org/10.2147/jmdh.s104807.
[22] Chowdhury R, Shah D, Payal AR. Healthy worker effect phenomenon: revisited with emphasis on statistical methods - a review. Indian J Occup Environ Med. 2017 Jan-Apr;21(1):2-8. https://doi.org/10.4103/ijoem.ijoem_53_16.
[23] Grimes DA, Schulz KF. Cohort studies: marching towards outcomes. The Lancet. 2002 Jan 26;359(9303):341-5. https://doi.org/10.1016/s0140-6736(02)07500-1.
[24] Faber J, Fonseca LM. How sample size influences research outcomes. Dental Press J Orthod. 2014 Jul-Aug;19(4):27-9. https://doi.org/10.1590/2176-9451.19.4.027-029.ebo.

[25] Kaplan RM, Chambers DA, Glasgow RE. Big data and large sample size: a cautionary note on the potential for bias. Clin Transl Sci. 2014 Aug;7(4):342-6. https://doi.org/10.1111/cts.12178.

[26] Konietschke F, Schwab K, Pauly M. Small sample sizes: A big data problem in high-dimensional data analysis. Stat Methods Med Res. 2021 Mar;30(3):687-701. https://doi.org / 10.1177/0962280220970228.

Article Information
Managing Editor: Jeremy Y. Ng
Peer Reviewers: Pallavi Dutta, Ya Ning Zhao, Curtis D'Hollander
Article Dates: Received Jul 19 22; Accepted Jan 31 23; Published Mar 30 23

Citation
Please cite this article as follows:
Zouhair J. Longitudinal prospective cohort studies: A research method primer. URNCST Journal. 2020 Mar 30: 7(3).
https://urncst.com/index.php/urncst/article/view/399
DOI Link: https://doi.org/10.26685/urncst.399

URNCST Journal
"Research in Earnest"

Funded by the Government of Canada | Canada