DOI: https://doi.org/10.4038/tsljos.v9i1.24

Selecting the appropriate study design for your research

Chandrarathne N[1], Senarath, U[1]

1 Department of Community Medicine, Faculty of Medicine, University of Colombo

Abstract

In the realm of clinical research, it is essential for both researchers and clinicians to possess a comprehensive understanding of different study designs, their inherent advantages, disadvantages, and the potential for bias associated with each.

Broadly, research designs fall into two categories: observational studies (e.g., cross-sectional, case-control, and cohort studies) and experimental studies (specifically, randomized controlled trials or RCTs). Each design serves a specific purpose, and they all come with their unique sets of strengths and weaknesses.

Selecting the most appropriate design hinges on two critical factors. Firstly, it depends on the specific research question at hand. For instance, if the research seeks to determine disease prevalence, a cross-sectional study is ideal; if it's centered around identifying harmful factors, a case-control study is more suitable; when examining prognosis, a cohort study comes into play; and for testing the effectiveness of a therapy, RCTs are often considered as the the gold standard. Secondly, practical considerations, including budget, time, availability of suitable patient populations, and the expertise of the research team, also play pivotal roles in determining the choice of study design. These factors can significantly constrain options.

While this paper provides a brief introduction to common study types, it is essential to recognize that conducting such studies requires access to far more comprehensive and detailed sources of information. As you embark on your research journey, you will likely need to delve deeper into these methodologies and seek guidance from more extensive resources.

Key words: study designs; research methodology; epidemiological studies; audit

1. Introduction

1.1. Background

Research is essential to understand the nature, causes and consequences of disease or health issues, and to provide the optimum solutions to address these problems. Well-designed research would generate evidence for health professionals to make right decisions regarding strategies for preventing illness, treatment options, and measures to minimize adverse outcomes on the affected individual, family and community at large. Health professionals engaged in patient care services are faced with several challenges in conducting research including inadequate knowledge in the research process, and adopting a technically sound methodology in a real-life clinical setting. Currently, there is a high demand, and increasing trend in conducting research in clinical settings under various disciplines. A simple and practical guide is essential for clinical researchers to plan and conduct a robust study with a greater potential for publication in a scientific journal, and for translation of evidence for practice and policy. This article guides researchers to select the appropriate epidemiological study design for a given research question.

■ Chandrarathne N. MBBS, MSc., MDS, MD, DFPH, FRSPH, Senior Lecturer, Department of Community Medicine, Faculty of Medicine, University of Colombo. nadeeka@commed.cmb.ac.lk

■ Senarath U. MBBS, MSc., MD, PhD, Chair Professor, Department of Community Medicine, Faculty of Medicine, University of Colombo. upul@commed.cmb.ac.lk

Correspondence: Nadeeka Chandrarathne, email: nadeeka@commed.cmb.ac.lk

Sri Lankan Journal of Orthopeadic 2023 : 08 : 40-46

1.2. Research, research problem and research question

Effective communication is vital in academic writing to ensure that both readers and reviewers can grasp your work and the underlying thought process with precision. Signposting, or providing clear markers of your research journey, aids in achieving this goal. The initial signpost is the purpose statement, which serves as the cornerstone for the entire study. Beginning with a broad and encompassing purpose statement, the researcher gradually narrows their focus. This narrowing leads to the formulation of specific questions that demand answers or predictions based on hypotheses that require rigorous testing.(1).

1.3. Objectives and hypothesis

In epidemiological studies, investigators use quantitative research questions and hypotheses, and more importantly objectives, to shape and specifically focus the purpose of their study. Quantitative research questions inquire about the relationships among variables that the investigator seeks to know (1). Quantitative hypotheses, on the other hand, are predictions the researcher makes about the expected relationships among variables. They are numeric estimates of population values based on data collected from samples. Testing of hypotheses employs statistical procedures in which the investigator draws inferences about the population from a study sample. Hypotheses are used often in experiments in which investigators compare groups. While hypotheses are usually used in formal research projects, such as a dissertation or thesis, objectives are used to indicate the goals or objectives for a study(1).

2. Study design

There are three main types of study designs: quantitative, qualitative and mixed methods. However, these should not be viewed as polar opposites or dichotomies; instead, they represent different ends on a continuum. A study tends to be more qualitative than quantitative or vice versa. Mixed methods research resides in the middle of this continuum because it incorporates elements of both qualitative and quantitative approaches. Often the distinction between qualitative and quantitative research is framed in terms of using words (qualitative) rather than numbers (quantitative), or using closed-ended questions (quantitative hypotheses) rather than open-ended questions (qualitative interview questions)(1).

2.1. Classification of study designs

Qualitative research is a means for exploring and understanding the meaning individuals or groups ascribe to a social or human problem. The process of research involves emerging questions and procedures. Data typically collected in the participant's setting. Data analysis - inductively building from particulars to general themes and the researcher making interpretations of the meaning of the data. The final written report has a flexible structure(1).

Quantitative research is a means for testing objective theories by examining the relationship among variables. These variables in turn, can be measured, typically by instruments, so that numbered data can be analyzed using statistical procedures. The final written report has a set structure consisting of introduction, literature and theory, methods, results, and discussion(1).

As most of the research done in surgical setups belongs to quantitative methods let us explore more on quantitative methods.

Before considering the pros and cons of these various study designs, it is important to review the fundamental concept of all quantitative clinical and epidemiologic research. Namely, is there an association between the exposure of interest (also known as: intervention, study factor, prognostic factor or risk factor) and the outcome of interest (disease or change in symptoms)? (Fig. 1). This, apparently simple relationship is complicated by the many other, additional risk factors (i.e. confounders or other exposures) that are also associated with that outcome (disease). For example, the subject’s age, gender, family history (proxy for genetics), socio-economic status and disease severity are all risk factors that can influence the outcome, and all will confound the primary association researchers are trying to establish. Irrespective of the chosen study design, a key issue is how will the researcher measure and effectively control for these confounding variables? Therefore, it is important to be aware of the inherent risk of bias with each of these study types, and therefore, the level of confidence clinicians will have in applying the results(2)

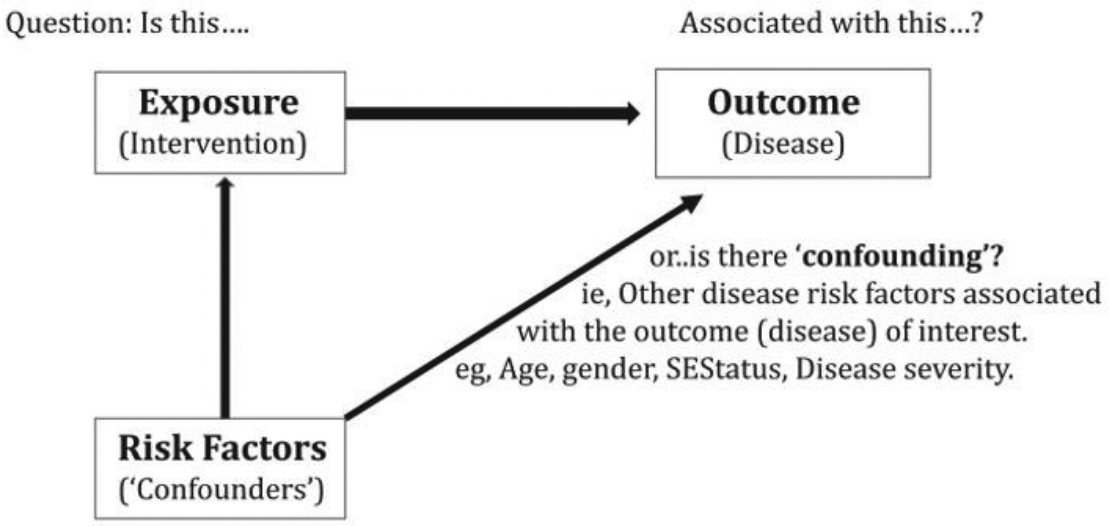

Figure 1: Relationship between exposure, outcome and confounders (adopted from Melis, 2020)(2)

Study designs can be broadly categorized as either observational or experimental. Observational means the researchers simply observe and record information about their subjects – but do not alter the study environment, while in experimental /interventional studies the researcher actively intervenes (experiments) the study groups (2).

Table 01 gives a summary of epidemiological study designs with the strength of evidence.

Table 1: The hierarchy of epidemiological study designs (adopted from Omir,2015 (3))

Study Design		Strength of evidence
Observational studies		
Descriptive Studies		
Case report	Single case	
Case series	Collection of similar cases	
Correlational studies	Population based study – using secondary data	
Cross sectional (descriptive)	Single sample from larger population – no comparison	
Analytical studies		
Cross sectional (analytical)	Single sample from larger population – compares two or more groups in the sample	
Case - control	Compares risk factors between diseased (cases) and non-diseased (controls) groups	
Cohort	Compares outcomes between groups exposed and non-exposed to a risk factor for a disease	
Interventional studies		
Clinical trials	Investigator allots the subjects to different groups - intervention versus non-intervention	

2.2. Observational studies

2.2.1. Descriptive studies

a) Case Report/Case Series

A case report is a presentation of an abnormal finding or outcome which, otherwise, would not be present. It describes the case in detail presenting the abnormal findings but is generally limited in demonstrating an association between the risk factor and the disease. The argument is that this finding may be due to chance or coincidence, without there being a real association between the risk factor and the outcome. Another example may be of an adverse (or beneficial) side-effect of a new drug, which may not have been documented before. These case reports are useful in identifying a new phenomenon and the reporting of similar cases from other observers, thereby helping generate a hypothesis for the association between exposure and outcome.

A "Case Series" is a collection of cases with the same outcome/finding. The general number of cases reported in a case series ranges from 20 to 50,but may vary from few as 5 to many as 100 or more. Case series are more useful than case reports in generating a viable hypothesis for an association between exposure and outcome. However, there is still a chance of this being only an incidental finding and the real reason may be due to some 'other' common factor. The main limitation of a case series study is the absence of a comparative group. So, it cannot be stated whether the outcome is really associated with the exposure or not, unless it can be shown that the group that is not diseased has a different exposure rate from the cases that are being studied. It is generally easy to select a matching group of subjects who are not diseased and determine the exposure to the risk factor in that group as well. This will convert the case series into an analytical case-control study design which is more useful in showing an association between the exposure and outcome. (3)

b) Correlational Studies

These are also called ecological surveys and are generally based on secondary data. They should not be confused with the term 'correlation' as used in statistical analysis. The term 'correlation' needs to be used with caution as it is misused/misinterpreted very commonly by researchers. The term 'association' is more appropriately used instead of correlation when establishing the relationship between different variables. Furthermore, it is important to remember that the correlation coefficient is used in statistical analysis to determine the association between two numerical (quantitative) variables. The use of the test for correlation can be applied in any type of analytical epidemiological study design, but this does not classify that study as a correlational study. Correlational studies, as described, use secondary data for two or more variables from different sources and try to determine an association between these variables. The most commonly used sources of secondary data are from governmental databases or reports of international agencies, e.g., cigarette sales per capita (in Rupees) versus incidence of lung cancer or number of cars registered in a city versus number of deaths due to road traffic accidents. This can be also done on an international level for different countries, e.g., literacy level or income per capita versus infant mortality rates. As discussed above, correlational studies are more commonly conducted at a national or international level. However, it is possible to conduct a correlational study on hospital-based data, e.g., number of patients being admitted in a day (from the hospital records) versus number of dosing errors or nosocomial infections . Correlational studies have the advantage of being quick and inexpensive since they are based on secondary data, which is already available from different sources. However, it should be kept in mind that any probable association shown in the correlational studies might be due to some other underlying factor and not due to the variables being considered. While there may be a link between a more sedentary lifestyle leading to increasing heart disease, the association may also be due to the difference in the gross income per capita between countries leading to difference in food consumption or higher cigarette consumption, etc.(3)

a) Cross-Sectional Studies

This is a relatively simple, quick, inexpensive study design used primarily to measure the 'burden of disease' (i.e. current prevalence) in a community, and to, ideally, identify potential risk factors for that disease. Because data collection on both exposure and outcome is usually based on questionnaire surveys, the scope for errors and biases with both measures is high.

The usual process is as follows: A representative sample from a community (population) is enrolled, and after informed consent, information obtained about their disease status or outcome of interest (e.g. 'doctor diagnosed asthma'; Yes/No), as well as information on 'exposure' measures. That is, possible 'risk factors' for asthma, such as atopic eczema and family history of asthma. Any association between these exposure variables and outcomes is then analysed, and expressed as an odds ratio (OR). The key design feature here is that both exposure and outcome are measured at the same point in time. Thus, there will be no loss to follow-up because participants are interviewed only once. A further strength is that numerous potential risk factors (exposures) can be measured in a single study. Despite some positives, cross-sectional studies are simply a 'snapshot', at a single moment in time, and are very prone to biased by estimates of exposures, outcomes and confounders. These studies are considered 'hypothesis generating', and possibly worthy of follow-up. As well as low-level evidence, these studies do not provide any information about cause-and effect relationships (2).

2.2.2. Analytical studies

a) Case–control study (CCS)

If the disease (outcome) is rare, or there is an urgent need for answers, a CCS may be the only practical choice in our set up. For example, the search for an explanation of a localised cluster of cancer cases. The basic design is to first choose disease cases (well-defined outcome) – for example cases of sudden infant death syndrome (SIDS). Each case is then matched with at least one control, of similar age, same gender and same community. While CCSs have an accurate outcome measure (case definition), the exposure measure may be problematic at times, because it is obtained retrospectively. In the stated example above it is, by interviewing parents of cases, and their respective controls. As expected, there is a high chance of errors, and particularly 'recall bias'. That is, cases and control parents will have a different likelihood of accurately recalling exposures, resulting in spurious associations. In addition, retrospective data on other risk factors (confounders) has the same high risk of errors and bias. For example, with the SIDS example, how accurately measured were other prognostic factors such as; maternal smoking, passive smoking, co-sleeping and socio-economic status. While it is easy to criticise; CCSs can sometimes hit the target, and can alter clinical practice. A classic paediatric CCS example was finding an association between SIDS and sleeping posture – leading to subsequent better evidence from prospective cohort studies. As with cross-sectional studies, given the inherent bias with CCSs, they are generally viewed as 'hypothesis generating', rather than influencing clinical practice. Depending upon the findings, ideally these studies should be followed up with a better study design, such as a cohort study. Clearly, there will be many obvious differences in risk factors between cases and controls (confounding) and statistical adjustment is inevitably incomplete – i.e. 'residual confounding'. As with cross-sectional studies, effect size is expressed as odds ratio (OR). In short, although quick, and relatively easy to perform, CCSs have a very high risk of finding erroneous associations, in both directions.

b) Cohort Studies

Prospective cohort studies

These studies recruit a representative sample of a population, such as very low birthweight infants (i.e. the exposure or study factor), who are then followed over a period of time, before determining the outcome of interest (e.g. learning difficulties at school entry). Along the way, any additional risk factors that arise which could affect the outcome are collected (e.g. seizures, hospitalisation), and controlled for in the analysis. As with CCSs, the strength of cohort studies is the ability to assess associations between multiple exposures and multiple outcomes. A further major strength of cohort studies ('prospective') is that researchers are able to accurately measure both the exposure of interest, and all known potential risk factors (confounders) at the outset of the study. Because additional risk factors are comprehensively measured, and controlled for statistically, the results of the cohort studies are more valid than CCSs. Because these studies generate a true incidence of disease, results are expressed as relative risk (RR). Again, as with all observational studies, residual confounding is still possible, thus downgrading the evidence from cohort studies, compared to a RCT. While cohort studies can more reliably identify associations than case– control and cross-sectional studies, they cannot establish a causal effect. While RCTs can overcome these prob-

lems, and supply higher level evidence, for practical and ethical reasons, they are often impossible. While cohort studies supply important evidence for paediatricians, there are some major drawbacks for researchers, including the long duration, expense, inevitable loss to follow-up and residual confounding.

Retrospective cohort studies

While a prospective design is usual, when accurate historical data concerning exposure is available, the study can be performed retrospectively. In this situation, the researcher begins the study after both the outcome and the exposure have already occurred. Clearly, this study design is only possible when the exposure of interest has been previously, accurately measured, recorded and stored in a secure database, such as a vaccination registry. Retrospectively, eligible subjects are then recruited, thus generating the cohort. Effectively, the follow-up observational period is 'historical'. Obvious advantages over the traditional prospective design, are speed, low cost, and ease of performance.

2.3. Interventional studies

In contrast to observational studies, randomised controlled trials (RCTs) enable accurate, reliable measures of exposure, outcome and randomisation effectively balances (controls for) all potential confounders, including unknown confounders. This assumes the randomisation is concealed, and the sample size is large. While RCTs are seen as the highest level of evidence, the proviso is that the study was internally valid, adequately powered, and well executed. As well as controlling for known and unknown confounders, RCTs can prove a true causal effect of an intervention (exposure). On the other hand, high-quality RCTs are very expensive, require very large on-going resources, generally take far longer than planned, and recruitment of adequate numbers of suitable subjects is always an issue, often delaying completion of the trial, or worse, publication of an underpowered RCT(2).

3. Clinical audit

A clinical audit is an exercise of measuring a clinical outcome or a process, against well-defined standards set on the principles of evidence-based medicine. Aim of the audit is to highlight the discrepancies between actual practice and standard in order to identify the changes needed to improve the quality of care. A peculiar characteristic of the clinical audit is the "professionalism" of the initiative, which is expressed by some typical ingredients: clinical specific competence of the participants, the confidentiality of the results, the object strongly connected to the "quality" of professionals (4). From a methodological point of view, clinical audit consists of a "quality loop"(Figure 2) : once chosen a topic and set shared and measurable criteria and standards, current clinical practice is evaluated, especially in terms of process or outcome, and suggestions for improvement are developed and applied, and then the cycle can begin again.

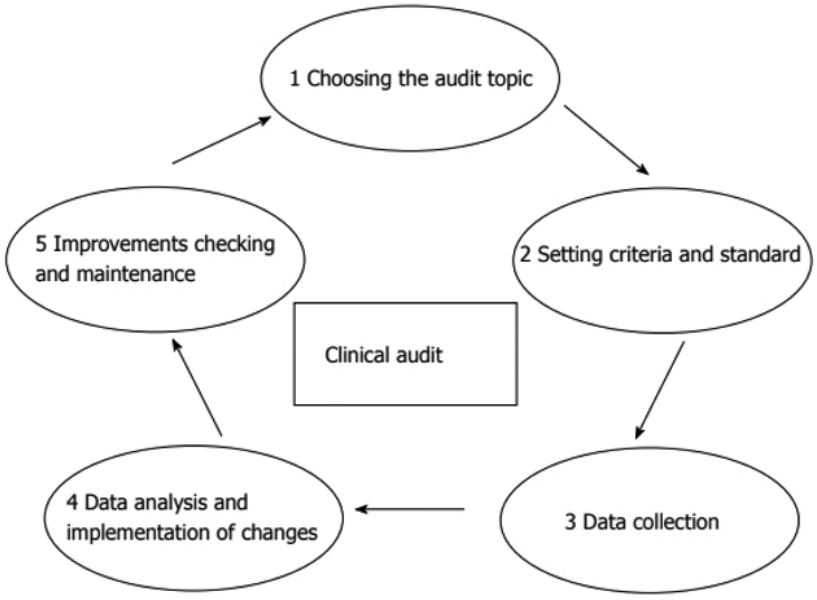

Figure 2: Clinical audit cycle

The audit should not be confused with data collection activities (i.e., benchmarking) or clinical research: the latter, in fact, aims to define the characteristics of good practice on an unknown land, while the audit compares the current practice against well-defined and established standards. The final aim of the clinical audit is always improving the care provided to the patient. As the 1992 BMJ article(5) clearly states "Research is concerned with discovering the right thing to do; audit with ensuring that it is done right" (5)

4. Approach to select the appropriate design

Given the many study design options, how does a researcher choose what type of study design? The first step (and one of the most difficult steps!) is to decide upon your research question. Once clarified, you will need to decide the optimal study design to answer that particular question. While this will be highly dependent on available resources, to a large extent the 'type' of question dictates selection of study type. For example, if the question is one of 'harm' – you will probably utilise a case– control study (CCS). If the question concerns 'prognosis/natural history/exploration of risk factors' – a prospective cohort study. And for a 'therapy' question – a randomised controlled trial.

The study types and their major purposes are listed in Table 2 (2).

Objective/Study purpose	Usual study type
Prevalence Disease burden	Cross sectional study
Risk factors	Case control/ cohort
Incidence/ New cases	Cohort
Prognosis Natural history Treatment effect To prove cause-effect	Randomized controlled trials

Table 2: Study types and their purpose

5. Conclusion

Quantitative clinical study designs come in various forms, each playing a crucial role in supplying evidence for the best clinical practices. These diverse designs come with their own set of strengths and limitations. It's important to note that observational studies, while sometimes criticized, often serve as the initial stepping stones toward more conclusive RCTs.

Furthermore, there are situations where conducting RCTs becomes either ethically impossible or practically impractical, especially in cases involving rare diseases or unique ethical considerations.

References

1. **Creswell JW.** Research Design: Qualitative, Quantitative, and Mixed Methods Approaches . 3rd ed. SAGE Publications Ltd.; 2009.
2. **Mellis CM.** How to choose your study design. J Paediatr Child Health. 2020 Jul 1;56(7):1018–22.
3. **Omair A.** Selecting the appropriate study design for your research: Descriptive study designs. Journal of Health Specialties. 2015;3(3):153.
4. **Esposito P**. Clinical audit, a valuable tool to improve quality of care: General methodology and applications in nephrology. World J Nephrol. 2014;3(4):249.
5. **Smith R.** Audit and research. BMJ. 1992 Oct 17;305(6859):905–6.