

https://dspace.org/

Considerations for choosing a research design

Patricia Mcinerney; Anushka Ajith; Deidre Pretorius; Sfiso Mabizela

https://hdl.handle.net/10539/46602
DHET

Downloaded from a DSpace repository. https://dspace.org/

Wits Journal of Clinical Medicine, 2024, 6(3) 1–4 http://dx.doi.org/10.18772/26180197.2024.v6n3a13

Statistics @ WJCM

Considerations for choosing a research design

Patricia McInerney*[1], Anushka Ajith[1], Deidre Pretorius[2], Sfiso Mabizela[3]

[1]Health Sciences Research Office, Faculty of Health Sciences, University of the Witwatersrand
[2]Department of Family Medicine, School of Clinical Medicine, Faculty of Health Sciences, University of the Witwatersrand
[3]Centre for Health Science Education, Faculty of Health Sciences, University of the Witwatersrand

***Corresponding author:** *Patricia.McInerney@wits.ac.za*

INTRODUCTION

Choosing a research design can be difficult, as it must align with the posed research question. The first decision is to select a research paradigm—either the quantitative paradigm, which is based on positivism, or the qualitative paradigm, which is based on interpretivism and constructivism. This manuscript will focus mainly on the quantitative paradigm, with a short note about qualitative and mixed-method research. Research designs in the quantitative paradigm are classified as experimental, analytical, and observational. Observational studies may be purely descriptive or analytical. A short description of these categories and an overview of some others will be provided, and important considerations will be discussed for each.

EXPERIMENTAL STUDIES

The most common form of experimental study used in the health sciences is the randomized control trial (RCT). Features of the experimental design are that the researcher manipulates at least one variable, usually the intervention, and that there is a control. Study subjects are allocated to either the intervention or control group through a process of randomization that aims to reduce potential bias. Randomization can be achieved using simple measures, such as tossing a coin, or by more sophisticated means, such as using random number generators. The importance of randomization and the formats it may take has been elaborated upon by Berger et al.(1)

An essential feature of RCTs is the use of 'blinding.' Blinding refers to a process whereby those involved in the study (participants, researchers, investigators, etc.) do not have information that could lead to bias, for example, by knowing whether a study participant was allocated to the treatment or control group. In some instances, blinding is not possible. Consider the case of an intervention to promote the active involvement of patients in their cancer care. A non-blinded trial was conducted to evaluate the efficacy of a web-based, personal patient profile-prostate decision aid versus the usual standard of care concerning decisional conflict in men with localized prostate cancer.(2) Participants would have known which 'intervention' they were receiving, and thus, it is clear that blinding would not have been possible in this study. It has been reported that researchers may have difficulty identifying what 'usual care' is when obtaining informed consent.(3)

Determining the sample size is critical to the study design and should be done during study development. This is important if the researchers wish to make inferences about the study results. Sample size calculators can be used and are easily found online, e.g., Calculator.net (https://Calculator.net>math) or Raosoft (http://raosoft.com>samplesize). While online sample size calculators can be convenient tools, it is essential to exercise caution as they often do not account for all the necessary parameters required for accurate sample size determination, particularly for complex study designs.(4) Consulting with a biostatistician early in the study design process is advisable to ensure accurate calculations tailored to the specific study context.

Randomised control trials manipulate at least one variable, so using a hypothesis is appropriate. A hypothesis is a prediction between variables. Rigby (5) states that statistical reporting in many RCTs is poor and provides a review of basic hypothesis testing. There are different types of RCTs, including cross-over trials, cluster RCTs, and stepped wedge trials. Cross-over trials are used when individuals in one study arm receive intervention A first and, after a while, intervention B. The individuals in the second arm of the study start with intervention B and, after a while, receive intervention A. Cluster RCTs are used when *groups* of individuals rather than individuals are randomised to treatment groups, and stepped wedge trials are used when an intervention is rolled out over a period of time. Copas et al. (6) identified three types of stepped wedge RCT designs and concluded that findings from these types of trials need to be reported more clearly. Pseudo-RCTs are used when randomization is not systematic, such as when individuals are allocated using odd/even numbers or days of the week.

Non-randomized control trials are an experimental design where participants are allocated to either the

article13.indd 1 03-10-2024 15:22:18

intervention or control group without randomization. These trials are used when randomization is impractical due to ethical or logistical constraints. To reduce selection bias and confounding factors, allocating participants to the intervention and control groups must be clearly defined and explained, and participants and researchers must be blinded to minimize bias. Propensity score matching can create comparable groups based on essential characteristics, and post hoc tests, including regression models, can be used to adjust for confounding variables.

Additional considerations for statistical analysis

When developing the proposal for a RCT, statistical analysis must be considered, mainly whether an intention-to-treat or per-protocol analysis will be conducted. The intention-to-treat method requires that the analysis is performed using the original numbers of participants allocated to each study arm, irrespective of whether they received or completed the treatment associated with the group they were initially assigned. Per protocol, on the other hand, uses the numbers in each arm that received the intervention at the end of the study, regardless of to which group they were assigned initially by randomisation.(7)

OBSERVATIONAL STUDIES

Observational analytical

This group of studies includes cohort and case-control studies. Song and Chung (8) note that RCTs are challenging to conduct in the surgical disciplines and that observational studies, such as cohort and case-control studies, can provide helpful evidence if well-designed. Cohort studies involve people who have a particular characteristic, who are then compared to people who do not have the characteristic. They may be studied either prospectively or retrospectively. In a prospective study, the exposure is measured before the outcome has occurred, while in a retrospective study, the outcome has already occurred and is linked to previous exposures. Prospective studies are more expensive to conduct.

Case-control studies are sometimes used to hypothesize about risk factors for a disease. Individuals in these studies have the disease or the outcome of interest. The comparative group does not have the outcome of interest. Cases and controls may be matched by age, sex, occupation, and other variables that may play a role in the outcome to control for confounding factors.

Observational descriptive

Cross-sectional studies, case series,(9) and case reports (10) are included in this group. Cross-sectional studies should select individuals in such a way that they are representative of the whole population so that the information generated can be extrapolated to the entire population. These studies help estimate the prevalence of a condition or characteristic, explore potential associations between variables, and assist in hypothesis generation. A limiting factor is that they cannot establish causality since exposure and outcome are measured simultaneously. When conducting a cross-sectional study, determining the sample size is necessary to have confidence in the results. One should make clear the population and, if appropriate, the target population and be able to estimate the population size, from which the sample size can be calculated. Kesmode (11) outlines the value of cross-sectional studies in obstetrics. The timing of data collection is vital to this design. For example, data may be collected over a relatively short period before a general election to predict the election results; on the other hand, data may be collected from the same individuals over a more extended period to determine changes in attitude towards a given concept. The latter type of data collection would constitute a longitudinal study.

Case series are studies in which only persons with a particular condition or disease are included. They are followed up on, and outcomes are reported. Such a design would be used when a new condition becomes evident in a community. A case report is similar but reports on one case only.

Record reviews

Record reviews usually take the form of retrospective research. Medical disciplines typically keep records of the types of diseases treated and details about the persons affected. Record reviews provide helpful information for planning future studies and allow for identifying trends over time. They have also been used in estimating adverse events.(12) Record reviews are an inexpensive research form, but they have shortcomings, such as incomplete records. Record reviews do not require a hypothesis because there is no variable manipulation, and the researcher works with the data available—a convenience sample.

Correlational research design

Correlation is a widely used and fundamental statistical concept in scientific research. Correlational research does not involve manipulating variables (13) but focuses on identifying and assessing relationships between variables. Two or more variables can be compared within the same population, or the same variable can be compared across different populations using correlational statistics.(14,15) The aim is to determine the direction of the relationship between the variables being compared. This relationship may be negative, positive, or exhibit no discernible pattern.

In correlational research, bivariate correlation is the most used parametric statistical analysis. It determines the correlation coefficient, accounting for the degree of a linear relationship between two variables. If the data is ordinal, a non-parametric statistical analysis, such as Spearman's Rho, is used.(16) However, depending on the research objectives, these statistical methods can be applied to other research designs.

A limitation of correlational designs is that while one can visually inspect the direction of the relationship between variables, this process is subjective, as is the case with scatter plots.(13) Additionally, the causes of the relationships cannot be determined. The strength of correlational studies, however, lies in their ability to help researchers understand not only the relationship but also the strength of the correlation between variables. For example, correlational research designs may use surveys to measure stress and academic performance. If such a study is conducted over more than one time point, it may constitute a longitudinal study.

Longitudinal research designs

Longitudinal studies involve an individual's continuous or repeated measurement over an extended period.(17–19) Typically observational, these studies can collect both qualitative and quantitative data. The key characteristic of longitudinal studies is their duration, often requiring a minimum of one year, and the observations must be consistent over time to assess changes effectively.(18,19) For example, a longitudinal study could be used to track the long-term effectiveness and safety of the COVID-19 vaccine in preventing infections and controlling the spread of the virus. Another longitudinal research might assess the impact of curriculum changes on students' academic performance and clinical practice outcomes.

Several previously described study designs, such as cross-sectional, prospective, and retrospective, can be used in longitudinal studies.(17) An example would be the electoral example presented earlier, where a cross-sectional study can be classified as longitudinal if the data is collected from the same individuals over an extended period to assess changes in attitude toward electoral candidates. Despite the variation in study designs that fall under the umbrella of longitudinal studies, the critical aspect remains their capacity to provide evidence of change within the studied population.(18) Like other research designs, longitudinal studies come with both advantages and disadvantages. A significant benefit is their ability to observe changes within a specific cohort over time, enabling researchers to establish the sequence of events.(17)

Additionally, longitudinal studies allow researchers to compare participants to their earlier selves, providing a more accurate description of the factors influencing the studied population.(18) However, these studies are often time-consuming and expensive due to the need for multiple assessments over an extended period. They also face challenges such as high attrition rates,(18) with many participants discontinuing their involvement and the potential for external influences to arise during the study, which can affect the results and generalisability.(20)

Qualitative research

Qualitative research seeks to interpret and understand a phenomenon of interest. There is no manipulation of variables. Data may be collected through various methods such as one-on-one interviews, focus group discussions, storytelling, and qualitative analysis of vignettes, documents, images, and audio or recorded material.(21) When conducting interviews, the researcher/interviewer uses a few pre-determined questions to guide the interview. Questions should be posed broadly to generate a detailed response, not monosyllabic answers. Responses may be probed through follow-up questions. The questions must remain faithful to the type of qualitative research being undertaken. Observation schedules can be open or structured according to pre-determined categories. There are different types of qualitative research, ranging from descriptive studies to grounded theory, phenomenology, ethnography, and critical theories. The latter group seeks to move away from interpretation only to the empowerment of individuals.

Data sufficiency is often questioned when assessing the trustworthiness of qualitative studies. Thus, it is important to note that sampling methods and the data collection structure can improve the study's rigor.(22) The neo-positivist notion also moves from data saturation to sufficiency.(23) Space does not allow detailed descriptions of the methodologies to be presented here, and novice qualitative researchers are advised to seek advice from an experienced qualitative researcher before embarking on such studies.

Mixed method research

Some research questions can be addressed through both the quantitative and qualitative paradigms. Consider the question of barriers and enablers to self-management in adolescents with asthma. This could be addressed through the quantitative paradigm by developing a structured questionnaire with both open and close-ended questions. It could also be addressed through the qualitative paradigm using one-on-one interviews or focus group discussions with adolescents with asthma. The researchers may want to use a mixed method approach here, collecting quantitative data using the questionnaire and then seeking to understand the numerical data through qualitative methods better. Researchers need to decide whether the quantitative or qualitative phase will be conducted first (sequential mixed method) or whether the two phases will be conducted simultaneously (concurrent). In the final phase of a mixed-method research design, the two phases must be combined and synthesized to provide a more complete and comprehensive picture of the phenomenon being studied.

CONCLUSION

Researchers are urged to spend time before starting a study considering the research question and making it clear and focused. The study design must align with the question to ensure that it is being addressed through the correct method. The novice researcher should discuss the study with a more experienced researcher whenever possible.

REFERENCES

1. Berger VW, Bour LJ, Carter K, et al. for the Randomization Innovative Design Scientific Working Group (2021) A roadmap to using randomization in clinical trials. BMC Med Res Methodol. 2021; 21(1):168.
2. Berry DL, Hong F, Blonquist TM, et al. Decision support with the personal patient profile-prostate: a Multi-Center Randomized Trial. J Urol. 2018; 199(1):89–97.
3. Nicholls SG, Carroll K, Zwarenstein, M. et al. The ethical challenges raised in the design and conduct of pragmatic trials: an interview study with key stakeholders. Trials 2019; 20:765.
4. Meysamie A, Taee F, Mohammadi-Vajari M-A, et al. Sample size calculation on web, can we rely on the results? J Med Stat Infor. 2014; 2:3. doi: 10.7243/2053-7662-2-3
5. Rigby A. Statistical methods in epidemiology: I. Statistical errors in hypothesis testing. Dis Rehab. 1998; 20(4):121–126.
6. Copas AJ, Lewis JJ, Thompson JA, et al. Designing a stepped wedge trial: three main designs, carry-over effects and randomisation approaches. Trials 2015; 16:352.
7. Andrade C. Intent-to-treat (ITT) vs completer or per-protocol analysis in randomized controlled trials. Indian J Psychol Med. 2022; 44(4):416–418.
8. Song JW and Chung KC. Observational studies: cohort and case–control studies. Plast Reconstr Surg. 2010;126:2234.
9. Din TF, Fagan JJ, Peer S profile of paediatric tuberculosis mastoiditis – a case series. S Afr J Surg. 2023; 60(1):62–66. https://doi.org/10.17159/2078-5151/2022/v60n1a3570
10. Koch GV. Carotid body tumour: a rare case report. S Afr Radiogr. 2022; 60(2):49–52.
11. Kesmode US. Cross-sectional studies – what are they good for? Acta Obstet Gynecol Scand. 2018; 97:388–393.
12. Thomas EJ, Studdert DM, Brennan TA. The reliability of medical record review for estimating adverse event rates. Ann Intern Med. 2002; 136:812–816.
13. Asamoah MK. Re-examination of the limitations associated with correlational research. J Educ Res Rev. 2014; 2(4):45–52.
14. Bloomfielde J, Fisher MJ. Quantitative research design. J Aust Rehabit Nurses Assoc. 2019; 22(2):27–30.
15. Curtis EA, Comiskey C, Dempsey O. Correlational research: importance and use in nursing and health research. Nurs Res. 2015; 6(1):20–25.
16. Allen P, Bennett K, Heritage B. SPSS statistics version 22: a practical guide. 3rd ed. Southbank: Centage Learning; 2014.
17. Caruana EJ, Roman M, Hernández-Sánchez J, Solli P. Longitudinal studies. J Thorac Dis. 2015; 7(11):E537–E540.
18. White RT, Arzi HJ. Longitudinal studies: designs, validity, practicality, and value. Res Sci Ed. 2005; 35(1):137–149.
19. Zhang W, Leng C, Tang CY. A joint modelling approach for longitudinal studies. J Roy Stat Soc Ser B: Stat Methodol. 2015; 77(1):219–238.
20. Harris MA, Orth U. The link between self-esteem and social relationships: a meta-analysis of longitudinal studies. J Pers Soc Psych. 2020; 119(6):1459–1477.
21. Patton M. Qualitative research and evaluation methods. 4th ed. Thousand Oaks, CA: Sage Publications; 2015.
22. Mwirta K. Factors influencing data saturation in qualitative studies. Int J Res Bus Soc Sci. 2022; 11(4):414–420.
23. Mays N, Pope C. Qualitative research: observational methods in health care settings BMJ 1995; 311:182–184.
24. O'Cathain A. Mixed methods research. In: Pope C, Mays N, ed. Qualitative research in health care. 4th ed. Oxford: Wiley Blackwell; 2020.

AQ1

AUTHOR QUERIES

AQ_1: Reference [24] is listed in the reference list but is not cited in the text. Please position the reference in the text in sequential order or delete it from the list.