UNITÉ DE SOUTIEN
SRAP | QUÉBEC

Composante Développements méthodologiques

Glossary of terms used in research

January 2020

McGill | Faculty of Medicine | Department of Family Medicine

Strategy for Patient-Oriented Research SPOR Putting Patients First

CIHR IRSC
Canadian Institutes of Health Research | Instituts de recherche en santé du Canada

UNITÉ DE SOUTIEN SRAP | QUÉBEC

Composante Développements méthodologiques
5858 Côte-des-Neiges Road, 3rd Floor, Suite 300
Montréal, Québec H3S 1Z1
Tel.: 514 399-9134 - supportunit.fammed@mcgill.ca - www.unitesoutiensrapqc.ca

Version (January 2020) prepared by: Christian Ruchon, Quan Nha Hong, & Paula Bush. Développements méthodologiques, Unité de soutien SRAP du Québec

How to cite: Ruchon, C., Hong, Q.N., & Bush, P. (2020, January). Glossary of terms used in research. Retrieved from https://osf.io/6r8wc/

If you have questions, contact: support.fammed@mcgill.ca - www.unitesoutiensrapqc.ca

UNITÉ DE SOUTIEN SRAP | QUÉBEC
Composante Développements méthodologiques
5858 Côte-des-Neiges Road, 3rd Floor, Suite 300
Montréal, Québec H3S 1Z1
Tel.: 514 399-9134 - supportunit.fammed@mcgill.ca - www.unitesoutiensrapqc.ca

Glossary of terms used in research

1. Absolute risk

It is the probability that an event will occur during a given period of time [10]. For example, a study includes two groups of 100 individuals each. One group will receive an intervention X (experimental group) and one group will not receive an intervention X (control group). The results of the study show that 10 individuals die after receiving intervention X while 25 die in the control group. The AR is 10% in the experimental group (10 ÷ 100) and 25% in the control group (25 ÷ 100).

2. Absolute risk reduction (ARR)

This is the absolute difference between the risk determined for the experimental group and that obtained for the control group as part of a study [10]. For example, with regard to the study on intervention X (see absolute risk), the ARR is: 0.25 - 0.10 = 0.15. This means that there is a 15% lower chance of death if a person receives intervention X compared to not receiving the intervention.

3. Case-control study

The case-control study aims to study the relationship between risk factors and a health problem. In this study design, two groups are formed : one group of individuals with a health problem under study (case) and another group whose individuals do not have the health problem (controls) and then the exposure to a risk factor is documented in both groups [6]. For example, a case-control study examined whether individuals with lung cancer (cases) are more exposed to smoking (risk factor) than individuals without lung cancer (control).

4. Case report

A report on a case (e.g., a person, a family) to describe in depth the different dimensions of the case [11].

5. Case series

A study design used to investigate several similar cases to describe in depth different dimensions under study [15]. A case can be defined as an individual, a family or an organization.

6. Case study

A study design whose purpose is a precise and clearly defined case based on specific parameters. A case can be defined as a person, an observed event or a document [1]. For example, a case study could study the behaviors of individuals in a department of a specific hospital between September 1, 2020 and May 1, 2021.

7. Code

A word or a sentence attributed to qualitative data that describes the general meaning of the data [5].

8. Cohort study

This study design consists in observing individuals and/or measuring certain outcomes without the researchers attempting to influence the intervention [10]. Cohort studies are a type of observational study in which a group of people exposed to a risk factor and a group of people not exposed to the same risk factor are monitored over time to determine whether exposure to that factor is likely to cause a health problem.

9. Conceptual framework

A conceptual framework is a configuration of an interrelated set of concepts, values and practices that provide a means of mapping the collection and analysis of data and the interpretation of results [3].

10. Confidence interval

In statistics, confidence intervals are used to estimate the expected level of confidence in the research results obtained. This is an interval around an estimated value where we are confident to include the value we are trying to estimate [12]. Within a confidence interval, there is a lower bound and an upper bound which are generally represented by brackets: [X - Y]. For example, a 95% confidence interval means that you can be 95% confident that the value you are trying to estimate is within that interval. More concretely, in a total population of 100,000 men, having a sample population of 10,000 men with an average weight of 180 pounds (95% CI [178 - 182]), means that you are 95% sure the average weight of all 100,000 men is between 178 and 182 pounds.

11. Correlation

The degree to which two variables seem to be associated. Correlation and causality should not be confused. Two correlated factors do not necessarily mean that one causes the other [6].

12. Cross-sectional study

A descriptive study in which a sample of people is selected and risk factors and health problems affecting this sample are measured simultaneously. Cross-sectional studies provide an overview of the health condition of the population at a given time. No association or cause-and-effect relationship can be concluded following such study design [6].

13. Data

Information collected to help answer the research question or test a hypothesis. This data can be in the form of numbers, words, images, sounds, etc.

14. Data analysis

A set of techniques and methods used to process the collected data. There is a variety of analytical methods depending on the nature of the data (qualitative, quantitative and mixed).

15. Deductive analysis

An analytical method guided by a conceptual framework or a theory already existing in the literature [1]. This type of analysis generally aims to test whether the data collected is consistent with the hypotheses or theories previously identified [2].

16. Descriptive values of screening tests

Screening tests are often used to identify who in a given population has a disease and therefore who does not. For technological, organizational, financial, or other reasons, these tests are not perfect. This imperfection is expressed through four situations that can occur following a test. The two most common are: a person is sick, and the test is positive (true positive) or a person is not sick and the test is negative (true negative). Although less common, it may also happen that a person has a disease and the test is negative (false negative) or that a person does not have a disease and the test is positive (false positive). Due to these false positives and false negatives, unique and specific values for each screening test allow

us to better describe the effectiveness of a screening test. These values are sensitivity, specificity, positive predictive value and negative predictive value.

17. Determinants
Risk factors that influence the presence of a health problem [6]. For example, poverty can be considered a determinant because it can influence a person's health.

18. Epidemiology
A field of study that allows studying in a population (a) the emergence, distribution and evolution of health problems, (b) any environmental circumstances or individual or social health-related conditions, and (c) disease control and health promotion [6].

19. Ethnography
A methodological approach that aims to document and describe a culture including social interactions, behaviors and perceptions that occur within groups, teams, organizations and communities. The central objective of ethnography is to provide a rich and comprehensive overview of people's views and actions through the collection of detailed observations and through individual and group interviews [3, 9].

20. Event rate
The event rate is the proportion of people in a group or population in which a particular event is observed [8].

21. External validity
The ability of a study to provide findings that can be generalized to another context and population. In other words, external validity describes the possibility that the results may be applicable to other populations and contexts.

22. Field observation
A method for collecting data that involves observing study participants in their natural context. These observations on the behavior and activities of individuals in this context, which constitutes the research site, are recorded in the form of field notes [5].

23. Focus group
A data collection method that consists of a semi-structured interview with a group of selected participants. A group is generally composed of six to twelve participants who share a common expertise or experience. There are usually a facilitator and an observer who takes notes. The facilitator's role is to facilitate interaction in the group and to bring out the participants' points of view, experiences, perceptions and beliefs.

24. Frequency
Corresponds to the number of events that occurred in a given group. For example, the number of people with diabetes in a population expressed as a proportion (e.g., 5%) [6].

25. Incidence
The measure of new cases occurring in a defined population for a defined period of time [6]. For example, the incidence of Lyme disease in 2018 was 329 in Quebec. In other words, 329 people

contracted Lyme disease in 2018, regardless of the number of people who previously had it.

26. Individual interview

A method of collecting data in the form of an interaction between an interviewer and an interviewee in order to gain an understanding of a phenomenon or problem [7]. The individual interview is a preferred means of obtaining information about the interviewee's knowledge, thoughts and experience on a specific topic. Interviews can be conducted face-to-face, by telephone or via the Internet.

27. Inductive analysis

An analytical method driven by the raw data rather than by a conceptual framework or a theory already existing in the literature [3]. This type of analysis allows raw data to be condensed and categories to emerge. It can also help to identify hypotheses or theories that will subsequently be explored in other studies [2].

28. Internal validity

The ability of a study to correctly estimate the object or effect that is intended to be estimated. This evaluation describes the extent to which the results are impartial and correspond to a precise estimate of the measured effect. The internal validity of a study depends on the design and implementation of a study. For example, this validity is increased if researchers have attempted to address all potential biases in a study [10].

29. Interpretation

The interpretation consists in explaining the significance of the study results [3].

30. Iteration

This involves repeating a specific process several times in order to have a better understanding of the data obtained [13]. More concretely, researchers review the data several times to ensure that the analyses and interpretations performed are accurate and clear. In practice, it is an ongoing process consisting in collecting data, conducting a preliminary analysis and using it to guide the next data collection and proceeding along this trend until data collection is completed.

31. Key informant

Key informants are individuals who help researchers gain and maintain access to the study environment and to better understand the data collected. The key informant serves as a guide, teacher and cultural facilitator for researchers [3].

32. Likelihood ratio

For results obtained in a given test (positive or negative), it is the ratio between the probability of being ill and the probability of not being ill [10]. This likelihood ratio is mainly used to assess the quality of a diagnostic test and to choose which test or sequence of tests is the most appropriate [16]. The interpretation of the likelihood ratio varies depending on whether the ratio is positive or negative.

A positive likelihood ratio is the probability that ill individuals will have a positive test result compared to the probability that individuals who are not ill will have a positive test result [10]. For example, a positive likelihood ratio of 5 means that there is five times more chance of a positive test result when the individual is ill than when the individual is not ill.

A negative likelihood ratio is the probability that sick individuals will obtain a negative test result

compared to the probability that individuals who are not ill will obtain a negative test result [10]. For example, a negative likelihood ratio of 0.20 means that there is five times more chance of having a negative test when the individual is not ill than when the individual is ill.

33. Mean

A measure obtained by dividing the sum of observed values by the number of observed values [6]. For example, in a set of values, 2, 2, 2, 2, 3, 4, 5, 5, 5 and 10, the mean is 4.

34. Median

The middle value that divides a set of ordered values into two equal parts [6]. For example, in a set of values, 2, 2, 2, 2, 3, 4, 5, 5, 5 and 10, the median is 3.5.

35. Mixed methods research

A type of research that combines qualitative and quantitative methods in the same study [17].

36. Mode

The most frequent value [6]. For example, in a set of values, 2, 2, 2, 2, 3, 4, 5, 5, 5 and 10, the mode is 2.

37. Negative predictive value

The probability of not having a specific disease when the value of the screening test is negative [10]. The higher this value, the more confident you can be that you will not have this disease if the test is negative.

38. Number needed to treat (NNT)

NNT describes the number of individuals to be treated in order to have a positive health-related event (e.g., an NNT of 15 means that for each group of 15 individuals treated with a specific drug, 1 heart attack is avoided). It is a measure of the effectiveness of a treatment so as to achieve a desired result. An effective treatment will have an NNT of 1. The higher the value of the NNT, the less effective the treatment is.

39. Odds ratio

A measure to calculate the association between an exposure to a potential cause and the occurrence of a disease in case-control study. This is an estimate of the relative risk.
How to interpret it?
OR = 1, there is no association between the exposure and the disease.
OR > 1, exposure to the risk factor increases the risk of having the disease (e.g., smoking increases the risk of having cancer).
OR < 1, exposure to the risk factor reduces the risk of having the disease (e.g., moderate alcohol consumption reduces the risk of cardiovascular disease).

40. P-value

A probability that indicates whether a result is statistically significant. For example, a value of "p <0.01" indicates that there is less than a 1% chance that the result is due to chance. The lower the p-value (if p<0.05), the more statistically significant the result is [6]. However, the p-value does not give an indication of the measurement, force or power of the effect.

41. Phenomenology
A methodological approach that aims to study the experience lived by individuals. Phenomenology can be defined as the in-depth investigation and description of the experiences or phenomena as experienced by the people who are living these experiences or phenomena [3].

42. Population
A set of specific elements (e. g., persons or objects) that have one or more common characteristics [6].

43. Positive predictive value
The probability of having a specific disease when the value of the screening test is positive [10]. The higher this value, the more confident you can be that you do have this disease if the test is positive.

44. Pre-test/post-test study or before-and-after study
It is a study design to evaluate the effect of an intervention (e.g., surgery, medication, therapy) in a group by measuring the same data with individuals before and after the intervention [8].

45. Precision
Precision is a measure that establishes how close the result(s) obtained from a study is (are) to the actual value. The scope of the confidence interval is considered as an indication of precision [14].

46. Prevalence
The proportion of people affected by a disease in a defined population within a defined period of time [6].

47. Qualitative research
Qualitative research focuses on how individuals and groups perceive, understand and conceptualize (e.g., understand and explain) phenomena [6].

48. Quantitative research
Quantitative research is generally used to test research hypotheses, evaluate the relationship between things or describe a population. This type of research makes it possible to quantify and measure by analyzing the data collected and by producing quantified results.

49. Randomized controlled trial
A type of research to study the effectiveness of an intervention. In this study design, people are randomly assigned to the experimental group (group that will receive the intervention under study) and the control group (group that will not receive the intervention under study) [8]. This random assignment is made without people being aware of the group in which they find themselves so as to reduce error or bias sources that could influence the results of experimental research. The purpose of random assignment is to make the different groups comparable.

50. Rate ratio
Ratio of risk in two different groups [6]. For example, one might be interested in comparing the level of risk of developing a given disease in a female population with that of a male population. A rate ratio can be calculated by dividing the level of risk in women by the level of risk in men.

51. Reflexivity

A process of critical self-reflection by a researcher on his or her prejudices, theoretical predispositions, preferences, etc., and their potential effects on the results of the study. This process allows the researcher to explore the effects of his or her presence on the context and social phenomenon he or she is trying to understand and to identify potential sources of bias [3].

52. Relative risk (RR)

An association measure that compares the risk of developing a disease in two different groups (a group A exposed to a given risk factor or to an experimental intervention and a group B not exposed to that risk factor or to that experimental intervention) [6].

How to interpret?

If RR = 1, the risk of developing a disease is the same in both groups;

If RR > 1, the risk of developing a disease is higher for group A than for group B;

If RR < 1, the risk of developing a disease is lower for group A than for group B.

53. Relative risk reduction (RRR)

The ratio resulting from calculating the difference between the probability of an event occurring in the experimental group and the probability of the event occurring in the control group divided by the probability of the event occurring in the control group. For example, in the study on intervention X (see absolute risk), the RRR is: (0.25 - 0.10) ÷ 0.25 = 0.6. This means that an individual who receives the intervention can expect a 60% reduction in the probability of death compared to an individual who does not receive it.

54. Research

Research is a rigorous and systematic process consisting of different stages of data collection and analysis aimed at a better understanding of a topic [6]. A research project consists of a research question and/or hypothesis, a research design, data collection and analysis.

55. Research design

Research designs are investigation models that provide specific guidance to the methods used in a research study [5].

56. Research hypothesis

A hypothesis is a statement that aims to explain or predict the expected relationship(s) between variables under study.

57. Research methods

The research methods correspond to the processes of data collection, analysis and interpretation of results [5].

58. Research question

A research question is a question about a specific problem that can be answered by a research project. The research question identifies the issue and concepts to be studied. It also influences the choice of research design, data collection methods and data analysis methods necessary to answer the question.

59. Results

This is the information obtained from the analysis of the data collected to answer the research question or hypothesis [18].

60. Risk
The probability that an event will occur (e. g., a disease) [6].

61. Sample
A subset of the population under study [6].

62. Sensitivity
The ability of a test to correctly identify individuals who actually have a disease [10]. In other words, it is the probability that a test result will be positive when the disease is present (true positive). A sensitive measurement tool prevents false negative results (false negative). If the sensitive test result is negative, a clinician is much more confident in excluding the possibility that a person may have a disease. A sensitive test will be sought if the consequences of a negative result are significant. For example, in the case of a serious disease or when there is an effective treatment for this disease, it is preferable to have a measurement tool with good sensitivity, which will prevent an individual with a disease from being undiagnosed.

63. Specificity
The ability of a test to correctly identify people who do not have a particular disease [10]. In other words, it is the probability that a test result will be negative when the disease is not present (true negative). A measurement tool with good specificity is preferable to confirm the presence of a disease since the risk of producing falsely positive (false positive) results is reduced. Highly specific tests are generally sought when it is important not to have a false positive. For example, when a treatment has major side effects (e.g., surgery, chemotherapy, etc.) it is preferable for the tool to have a good specificity since it prevents a person from being misdiagnosed.

64. Standard deviation
A value indicating the dispersion of the data around the mean [6].

65. Statistical inference
Method used to draw conclusions on a population of interest based on results obtained in a sample.

66. Statistical power
Statistical power describes the ability of a study to detect a difference between two groups. The power of a study depends on several factors such as the sample size, the reliability of measurement instruments, and the appropriateness of statistical tests [15].

67. Statistically significant
A statistically significant result means there is a high probability that the difference between the results is due to something other than chance. In contrast, when the results of a study are not considered statistically significant, it means that the results could have occurred by chance. However, this does not mean that there is no effect (e.g., it cannot be concluded that the intervention does not reduce pain), but rather that the study does not provide sufficiently convincing evidence of the effect. This situation may occur, for example, when the sample size is too small [10].

68. Subjects
The study objects, entities, etc., for which the values of the variables are measured (often humans, but may also be animals or living cells in culture, among others) [6].

69. Thematic analysis
An analytical method that aims to identify and analyze themes that emerge from a set of qualitative data [4].

70. Theme
It is about ideas or topics that come from the data. The themes capture the meaning and scope of the meaning of a specific group of data. They can be used to group data that might otherwise appear distinct, or to group meanings that occur in multiple and varied contexts [4, 9].

71. Transferability
A term used in qualitative research to address the potential applicability of the results to other subjects or contexts [1]. This term is similar to the term external validity or generalizability in quantitative research.

72. Triangulation
Refers to the use of several theories, methods or data sources on the same subject to increase the reliability of the results and develop a more complete understanding of the phenomenon under study [1, 19]. For example, to better understand the interaction between a physician and a patient, a researcher could use questionnaires, observations and interviews as a data source rather than just one of these sources.

73. Validity
Ability of a study or measurement instrument to correctly understand a phenomenon [12]. For example, the validity of a screening test is the ability of a test to properly identify individuals who have a disease and those who do not [15].

74. Variables
Characteristics that are measured or documented (e.g., body weight, age, sex, blood cholesterol levels, smoking, etc.) [6].

Composante Développements méthodologiques
5858 Côte-des-Neiges Road, 3rd Floor, Suite 300
Montréal, Québec H3S 1Z1
Tel.: 514 399-9134 - supportunit.fammed@mcgill.ca - www.unitesoutiensrapqc.ca

References

1. Green, J. and N. Thorogood, *Qualitative methods for health research*. Third ed. Introducing qualitative methods. 2014, Los Angeles: SAGE.
2. Blais, M. and S. Martineau, *L'analyse inductive générale: description d'une démarche visant à donner un sens à des données brutes.* Recherches qualitatives, 2006. **26**(2): p. 1-18.
3. Schwandt, T.A., *The Sage Dictionary of Qualitative Inquiry*. 2015, Thousand Oaks, CA: SAGE Publications.
4. Braun, V., et al., *Thematic analysis*, in *Handbook of Research Methods in Health Social Sciences*, P. Liamputtong, Editor. 2019, Springer: Singapore. p. 843-860.
5. Creswell, J.W. and J.D. Creswell, *Research design: Qualitative, quantitative, and mixed methods approaches*. 2017, Thousand Oaks: SAGE.
6. El Sherif, R., et al., *Atelier sur la littératie en recherche: Cahier du participant.* McGill Family Medicine Studies Online, 2018. **13**(e02).
7. Baribeau, C., et al., *L'entretien individuel en recherche qualitative : usages et modes de présentation.* Revue des sciences de l'éducation, 2012. **38**(1): p. 23-45.
8. Gouvernement du Canada. *TERMIUM Plus®*. 2019 [cited 2019 25 juillet]; Available from: https://www.btb.termiumplus.gc.ca/.
9. Reeves, S., A. Kuper, and B.D. Hodges, *Qualitative research methodologies: ethnography.* BMJ, 2008. **337**: p. a1020.
10. BMJ Best Practice. *A glossary of EBM terms*. 2019 [cited 2019 25 juillet]; Available from: https://bestpractice.bmj.com/info/us/toolkit/ebm-tools/a-glossary-of-ebm-terms/.
11. Gustafsson, J. *Single case studies vs. multiple case studies: A comparative study*. 2017.
12. Simpson, A., C. Beaucage, and Y. Bonnier Viger, *Épidémiologie appliquée: Une initiation à la lecture critique de la littérature en sciences de la santé.* 2009, Montréal: Gaëtan Morin Éditeur, Chenelière Éducation.
13. Srivastava, P. and N. Hopwood, *A practical iterative framework for qualitative data analysis.* International journal of qualitative methods, 2009. **8**(1): p. 76-84.
14. Wade, A., *Power and precision in research.* Archives of Disease in Childhood, 2018. **103**(3): p. 280.
15. Gordis, L., *Epidemiology*. Fifth ed. 2014, Philadelphia, PA: Elsevier/Saunders.
16. Centre for Evidence-Based Medicine. *Likelihood Ratios*. 2019 [cited 2019 25 juillet]; Available from: https://www.cebm.net/2014/02/likelihood-ratios/.
17. Bujold, M., et al., *Cahier scientifique Acfas #117 - Oser les défis des méthodes mixtes en sciences sociales et sciences de la santé*. 2018, Acfas: Montréal.
18. Fortin, M.F., J. Côté, and F. Filion, *Fondements et étapes du processus de recherche*. 2006, Montréal: Chenelière Éducation.
19. Carter, N., et al., *The use of triangulation in qualitative research.* Oncology Nursing Forum, 2014. **41**(5): p. 545-7.