6-Part Webinar Series: Research Methodology
Part 1: Study Types and Risk Estimates

Questions Asked: January 18, 2024

1. **What is the upper limit for the number of participants in a case series?**

   As we discussed – it really depends on the topic and the number of cases relevant to the specific topic. A description of the different cases with regards to similarities and differences would be appropriate.

1. **As field epidemiologists, especially for outbreak investigation, what would be the best study design to select?**

   Again – dependent on the outbreak (severity), amount of time, and available resources and data. Thus, could be a cohort study, case-control or cross-sectional approaches.

2. **Journals that do not like case reports/series are limited to non-communicable diseases. Please share your thoughts.**

   Certainly important to check the journal specs to assure case reports/series are included. Also check specific cases that are published to assure your submission is appropriate for the journal.

   For the most part, every specialty values and publishes case reports. However different specialties will have different interests in what a case report should include. For example, a case report for a endocrinology journal will have a different focus than if it was written for a neurosurgery journal. For the endocrinology, readership and editors might find details related to laboratory findings to be more interesting than surgical pictures and vice versa.

3. **What is the difference between a case study and a case report?**

   Pretty much the same – though a case study can include multiple cases.

4. **What is the minimum number of cases in a case series? What/when is a good time to report the case series?**

   The number of cases depends on the topic and relevance to the area of study. The timing is dependent on the relevance of the case and contribution to practice.

5. **What is the difference between a time series or case series and a step-wise time (perhaps stepped-wedge)?**

   Time series analysis is a specific analysis of a sequence of data points collected over an interval of time.

   In stepped-wedge designs (most often considered for clinical trials) , the clusters are randomized into several groups or waves that define when the intervention will begin, and all clusters start the trial in the control condition. Groups of clusters cross over to the intervention condition on a staggered schedule.

6. **How good is using the Case series in a systematic or scoping review?**

   Typically not included – but with specific conditions or stage of research, case reports can certainly be included in a search when this is the major source of evidence.

7. **If I understand well, are retrospective cohort studies abnormal?**

   Apologies for any confusion – retrospective (historical) cohort studies can be very valuable WHEN good records and information is available on the cohort. It might be best to refer as a historical cohort.

8. **I would like to understand more about the controversy that retrospective is the same as prospective, as the professor said.**

   A retrospective (historical) cohort studies provides valuable results but needs high quality data and health indicators, such as detailed health records such to be able to calculate incidence. The "controversy' regards how the cohort is referred – historical is perhaps a better term.

9. **Is it not possible to have a retrospective cohort study? Can you clarify the direction of inquiry? Can we not compute the incidence of a disease that occurred in the past?**

   Certainly retrospective (historical) cohort studies have value and possible to calculate incidence WHEN adequate details records are part.

10. **In terms of terminology, what suggestions do you have for defining the term 'retrospective cohort' while maintaining a clear distinction from 'prospective cohort'?**

    Prospective cohort – following a cohort when the event has NOT occurred.

    Retrospective (historical) cohort – event has occurred and factors before the event are assessed with historical records and data.

11. **Please give one example of a retrospective Cohort.**

    In a retrospective (historical) cohort, investigators use preexisting data to identify exposed and unexposed individuals in the past, without regard to outcome status, and trace these individuals forward, up to and possibly including the present, to determine incident outcomes.

12. **Is it also possible to do a case-control study using secondary data like a retrospective cohort study design?**

    While both use historical data – the cohort study follows a group and estimates incidence and relative risk. A case-control study is not able to determine incidence and thus calculates an odds ratio.

13. Can you explain why cohort studies are prospective?

Can be either historical or prospective. Prospective is typical as allows assessment of exposure.

14. Is it possible to have a purely descriptive cohort with no exposed or unexposed groups?

Yes, depends on critical question being answered.

15. How do we name a study when a study has case and control groups, yet not asking for a risk factor of the disease of the case group, but measuring and comparing the current situation of both groups?

Could be considering other factors to explain differences between cases and controls.

16. What will be the difference between a prospective study, which takes years, and a longitudinal study?

Very similar with defined cohort.

17. What are your thoughts on nested case-control studies?

Valuable - In the nested case-control study, cases of a disease that occur in a defined cohort are identified and, for each, a specified number of matched controls is selected from among those in the cohort who have not developed the disease by the time of disease occurrence in the case.

18. What about Nested case-control based on cohort? Could it be used to gather the advantages of both case-control and cohort studies?

The main difference between a case-control and a nested case-control study is that in the former the cases and controls are sampled from a source population with unknown size, whereas the latter is 'nested' in an existing predefined source population with known sample size.

19. Could you say something about nested case-control study design? Why would one choose it instead of a normal case-control study?

The main difference between a case-control and a nested case-control study is that in the former the cases and controls are sampled from a source population with unknown size, whereas the latter is 'nested' in an existing predefined source population with known sample size.

20. I would like your comments and recommendations about the critical moment of selecting cases and control, which I understand is crucial to the study quality.

As discussed matching controls is critical for a case-control study

21. Can you do a study and include both designs? Can we conduct a cohort and a case-control design in one study?

In the nested case-control study, cases of a disease that occur in a defined cohort are identified and, for each, a specified number of matched controls is selected from among those in the cohort who have not developed the disease by the time of disease occurrence in the case.

22. Can you please explain the type of design that combines different study types, such as prospective case-control, cohort, crossover, etc.?

Please see above.

23. For a case-control study: is it possible to recruit case and control from different diagnosis times?

Possible but you want the cases and controls to match as close as possible.

24. Can a retrospective cohort study be used to analyze NCD risk factors (e.g., diabetes) using NCD risk factor screening data (secondary data)? The data contains names and risk factors that are only measured once.

Yes, if adequate records are available.

25. Why is a case-control study referred to as quick and dirty?

Able to be completed in a short period of time.

26. Which kind of study is more valuable than the other: cohort study or case study?

Cohort study allows estimate of incidence.

27. Why do the interpretation of case-control and cohort study findings are similar even though their objectives are different? Is that appropriate to interpret the risk of occurrence from a case-control study (Why not exposure differences)?

Cohort is used with incidence.

Case-control – odds ratio

28. Is there a specific rule on interpreting the computed ratio (either odds or risk)? For example, if we computed an odds ratio of 1.5, do we put 2 in the interpretation or leave it as is?

Typically – the calculated 1.5 would be reported.

29. What's the difference between 'risk difference' and 'odds ratio'?

The risk difference (RD) is the difference between the absolute risks of 2 interventions or risk factors. The RD represents excess risk attributed to the group with the higher risk. The odds ratio can estimate the risk ratio.

30. Is there really a clear delineation between the questions answered by odds and risk ratios? Or is it just a matter of the direction of inquiry?

The relative risk (also known as risk ratio [RR]) is the ratio of risk of an event in one group (e.g., exposed group) versus the risk of the event in the other group (e.g., nonexposed group). The odds ratio (OR) is the ratio of odds of an event in one group versus the odds of the event in the other group.

31. Is there a case/situation where the odds ratio approximates or equals the risk ratio?

The relative risk (also known as risk ratio [RR]) is the ratio of risk of an event in one group (e.g., exposed group) versus the risk of the event in the other group (e.g., nonexposed group). The odds ratio (OR) is the ratio of odds of an event in one group versus the odds of the event in the other group.

32. If the odds ratio approximates relative risk, how can we know up to what percentage?

Relative risk from a cohort study includes incidence and is the better predictor.

33. What will be our basis in choosing between odds and risk ratio for cohort studies?

The relative risk (also known as risk ratio [RR]) is the ratio of risk of an event in one group (e.g., exposed group) versus the risk of the event in the other group (e.g., nonexposed group). The odds ratio (OR) is the ratio of odds of an event in one group versus the odds of the event in the other group.

34. You said we use case-control for the odds ratio. What study design would I use for incidence? Would it be cohort?

Correct – cohort.

35. RR is appropriate for cohort study, but shall we use RR for historical cohort study? we have faced difficulty during our analysis? AOR?

Relative risk is appropriate.

36. Can you please explain the difference between absolute and relative risks?

Absolute risk refers to the actual probability of an outcome occurring in a specific group regardless of any other factors. In this case, it would be the chance of being struck by lightning. Relative risk on the other hand, compares the risk of an outcome between exposed and unexposed groups.

37. What is the difference between odds of ratio vs prevalence ratio?

Odds ratio (OR) and risk ratio (RR) are two commonly used measures of association reported in research studies. In cross-sectional studies, the odds ratio is also referred to as the prevalence odds ratio (POR) when prevalent cases are included, and, instead of the RR, the prevalence ratio (PR) is calculated.

38. I read cross-sectional studies that compute the odds ratio. Can you provide a brief discussion on this? Can we also compute the risk ratio in cross-sectional studies?

Odds ratio (OR) and risk ratio (RR) are two commonly used measures of association reported in research studies. In cross-sectional studies, the odds ratio is also referred to as the prevalence odds ratio (POR) when prevalent cases are included, and, instead of the RR, the prevalence ratio (PR) is calculated.

39. In a combined study design, what would be the measure of risk?

Depends if incidence can be determined.

40. Can we calculate relative risk for case-control studies?

Odds ratios are determined from case-control studies.

41. How can we interpret the study results with RR=3.3 and Confidence Interval=0.8-4.9?

As 1 is in the CI – the risk estimate would not be considered statistically significant.

42. Can we also calculate the hazard ratio of the disease if we do a case-control within the cohort studies?

In a study that is designed and conducted as a case-control study, you cannot calculate incidence. Therefore, you cannot calculate risk ratio or risk difference. You can only calculate an odds ratio. However, in certain situations a case-control study is the only feasible study design.

43. Describe a situation when calculating the hazard ratio is better than the risk ratio.

One of the main differences between risk ratio and hazard ratio is that risk ratio does not care about the timing of the event but only about the occurrence of the event by the end of the study.

44. Why, in this study, did they estimate the HR and not the RR? (slide 39)

One of the main differences between risk ratio and hazard ratio is that risk ratio does not care about the timing of the event but only about the occurrence of the event by the end of the study

45. Can you say a word on the use of HR in cohort studies not studying survival?

A measure of how often a particular event happens in one group compared to how often it happens in another group, over time. Does not have to be death.

46. Professor Whelton pointed out the best estimate… what confidence can you have in that number… Isn't the Confidence Interval the tool to answer this?

The confidence interval addresses precision which can be a part of 'confidence'.

47. How do I determine if a confidence interval contains one and that it is significant?

There are occasions when a statistical test provides a different statistical significance from the CI.

48. How do you differentiate that Retrospective Study on breast cancer, based on routine information collection, from a case series?

This was a historical cohort as opposed to a group of cases.

49. In the case-control study for breast cancer exposure to the pesticide, how were the controls agreed on and selected? (slide 24)

Cases (N = 155) were recruited from a population-based cancer registry, and controls (N = 150) were obtained from tax assessor and Medicare list mailings.

50. The example you are explaining is it not a nested-case control study design since the controls are obtained from a population-cohort ?(slide 25)

Sorry for any confusion – this was presented as a case-control study.

51. Can the Ischemic stroke study be a nested case instead? (slide 25)

Certainly ischemic stroke could be assessed in a nested case control study – but not the case.

52. How could we explain the wide confidence interval despite the range does not include 1? since it is based on a population-based cohort (slide 25)
Atrial fibrillation (OR, 10.43; 95% CI, 2.33–46.77], CVD (OR, 8.01; 95% CI, 3.09–20.78), type 2 diabetes mellitus (OR, 2.31; 95% CI, 1.35–3.95), LDL cholesterol (OR, 1.81; 95% CI, 1.37–2.40), current smoker (OR, 1.81; 95% CI, 1.50–2.17), hypertension (OR, 1.43; 95% CI, 1.17–1.75), family history-stroke (OR, 1.37; 95% CI, 1.04–1.82)

You are correct – the wide ranges are complicated to explain.

53. **What is the cut-off point to say a confidence interval is wide or narrow?**

    Depends on the critical question and issues being addressed.

54. **If the number of controls is smaller than the number of cases, does this create bias in the study?**

    Depending on the topic – a minimum of 1:1 match with typical 2-4 controls per case.

55. **Can we accept a case-control study in which the number of controls is less than four times the number of cases?**

    Depending on the topic – a minimum of 1:1 match with typical 2-4 controls per case.

56. **Can we estimate risk factors for HTN in the community using a case-control design measuring exposure and outcome at the same time?**

    Yes, but include appropriate valid parameters.

57. **In research on preeclampsia and eclampsia, is case-control a good design when doing factors that could increase preeclampsia?**

    A case-control can be a useful design for preeclampsia.

58. **What study design would I use to determine the incidence of…?**

    Typically a cohort design.

59. **For instance, the clinics and epidemiology of thalassemia patients in Afghanistan in 2023; is it a cross-sectional, case series, or incidence study?**

    It would depend on the critical question being addresses, but any if these designs could be employed.

60. When we describe the epidemiology of cases for one year in the hospital, what design is this? Sometimes, it is said to be a chart or record review; is it also a case series?

Could be a cross sectional or case series.

61. Can we have a short-term cohort study? As in, no follow-up?

There couple be a passive follow-up where you follow the cohort till death or event as documented by records- but no active follow-up or contact with the individual.

62. Is there a clear-cut point for rare diseases/ exposure? How much of the population?

A rare disease is based on how prevalent it is—that is, the number of individuals living with it.

In general - In the United States, a rare disease is one that fewer than 200,000 people live with. (In other words, 60 per 100,000 individuals.) Around the world, rare diseases are identified and addressed differently. The European Union considers a disease rare if it affects no more than 50 per 100,000 people.

63. Using a retrospective study, can we examine a temporal relationship between a risk factor and an outcome of interest?

Yes.

64. In a retrospective study that relies on an available clinical database, how do you select the appropriate inclusion criteria and choose your data element to request from the database owner?

Basically the same approach as a prospective study except dependence on existing parameters from various records and data bases.

65. If one conducts an online survey, where does this lie in the evidence pyramid?

If part of a cross-section study-observational study.

66. Can census data be used to answer epidemiologic questions?

Valuable for population demographics and changes in population.

67. Is there a way of determining the exact strength of the association between exposure and outcome?

Strength of an association can be assessed by considering the incidence of an event in an identifiable group (numerator) and comparing that with the incidence in a baseline group (denominator). A relative risk of 1 indicates no association, whereas a relative risk other than 1 indicates an association.

68. In epidemiological studies, researchers often face the challenge of establishing causation versus correlation. Can you propose a methodological approach that effectively addresses confounding variables and strengthens causal inference in observational studies?

Correlation means there is a statistical association between variables. A correlation between variables, however, does not automatically mean that the change in one variable is the cause of the change in the values of the other variable. Causation indicates that one event is the result of the occurrence of the other event; i.e. there is a causal relationship between the two events. Basically, causation means that a change in one variable causes a change in another variable.

69. Can an observational study be run with control parameters collected from literature and not on the field?

Typically a study would include collected data for the control. However, when controls are difficult or impractical to obtain, historical controls can be used for model parameter estimation at the study design phase, adaptation within a study, or supplementation or replacement of a control arm. Real-world data is one of the primary resources for constructing historical control datasets, and include patient registries, medical charts, and systematic reviews of published clinical data. Regardless of the source, it is imperative to accurately and reliably capture data at clinically relevant time intervals.

70. When estimating risks in clinical trials or public health interventions, how can one reconcile the tension between the need for statistical significance and the ethical implications of potentially exposing participants to risks or withholding potentially beneficial treatments?

Ethical considerations and human subject protection are indeed the most important issue in a clinical trial. The 4 main ethical principles are beneficence, nonmaleficence, autonomy, and justice, are defined and explained.

71. Can a person conduct his study and have his data by only using a questionnaire? Will this allow him to restrict and take care of his control group?

It does depend on the specific topic and study, but questionnaires can have great value to a study.

72. Could you please present more examples of food safety?

Yes indeed.

73. Could it be possible to take examples from animal diseases also?

We will try to include but will continue to focus on human disease. There are some good examples of both human and animals.

74. Please present more examples of high-quality global research methodology, e.g., global methodology.

Will do.

75. If I want to calculate the prevalence of HIV in a specific population, and among the participants, I have those who were tested during the study and those who reported living with HIV (previously tested). Since prevalence is looking at existing cases, will my numerator include those who reported living with HIV?

Yes.

76. What study would be helpful in conducting a research study for Mental health illness (psychiatric) disease when we make the diagnosis/prognosis only through clinical signs and symptoms (without any laboratory tests)? What type of study would be helpful to deduct the etiology?

It might be useful to begin with the cross-sectional observation study.

77. What study designs can be done for hospital infection prevention and control studies?

Such benefit could be determined from a case series, case-control, or cohort study depending on the available resources and time.

78. **Which kinds of studies besides case-control studies can we use to get strong evidence on the association between some factors and the given disease?**

    There is value from most any study including case series, case-control, and cohort study, each design with different evidence levels. All is driven by the available resources and time.

79. **Could you please explain and address the advantages and disadvantages of ecological studies?**

    Ecological studies are typically considered relatively cheap and easy to perform by using population data that has already been collected and seeking correlations between potential risk factors and various disease outcomes. However, data are unlikely to be collected for all members of the group of people considered as the unit of analysis. Thus, when the data are aggregated across the unit of analysis in an ecological study, the outcome measures are likely to be inadequate or biased.

Logistics

1. **Are those questions and answers raised here going to be organized with slides? If so, even better?**
   Yes, and we will do this after each session.
2. **I am from Peru; please do not accept the subtitled translation in Spanish.**
   Thank you and noted.
3. **Sorry, the presenter is too fast.**
   Will try to slow the pace.
4. **Last time, you promised to provide an interpreter system in French, but I can't see anything allowing me to activate interpretation.**
   Sorry for any confusion but the current sessions do not include interpretations.
5. **Is it possible to disable chat for now? The introductions are distracting.**
   Noted.