Study design II – Practical study design

Gustaf Edgren, MD PhD
Associate professor of epidemiology
Karolinska Institutet

Material covered in CW 1

- In Epi I you went into some detail for the two general types of studies:
  - Observational studies:
    - Cohort studies
    - Case-control studies
    - Etc.
  - Intervention studies
- We did, however, not cover details on how such studies are actually conducted…

The epidemiologic research process

- Generally, the conduct of most studies follows a similar process:
  - Study concept/idea/hypothesis
  - Reading up on background for studies
  - Formulation of hypothesis
  - Planning of study to answer this hypothesis
  - ~~(Acquiring funding to conduct study)~~
  - Practical conduct of study
  - ~~Analysis~~
  - ~~Interpretation~~
  - ~~Publication~~

Study concept/idea

- Often, ideas for research comes to us from unexpected sources
- It might be from some clinical observation, some random thought, some novel biological finding, or from reading someone else's work

Study concept/idea

- I've noted that once you start thinking about clinical problems and questions from an epi perspective, you'll see important studies everywhere
- Tip: Keep a list of your ideas!
- You'll notice, however, that once you do the due diligence, most of your ideas will fail (or will already have been tried by other researchers)

Study concept/idea

- Once you've actually found an idea for a study that has stuck with you, the following is in place:
  - Has anyone else already gotten this idea? (Check PubMed and Google Scholar)*
  - Can you immediately think of some way this idea could be tested? (Don't be too discouraged if you can't)
  - Ask yourself if the best possible result of the study would be meaningful and interesting
  - Tell a friend about your idea (I generally send everything to Mikael)

Study concept/idea

- If your idea persists (i.e. it hasn't been done thoroughly enough before, its testable, the results would be interesting, and your friend thinks it's a good idea), do a more thorough background check of the literature*
- At this stage, it is probably also wise to get smart about specifics on how to design the study:
  - What is the natural design?
  - Where can I get the data?
  - What are the possible limitations?
  - Who do need to involve (and ask permission from)?
  - Etc.

Background check

- I assume you've all done literature searches, but here are some general tips:
  - PubMed is often far too overwhelming so start elsewhere (e.g. Google Scholar) or limit yourself (e.g. core clinical journals, reviews)
  - Don't forget the old studies that are often not found in PubMed – research didn't start with the advent of computers!
  - Scan the reference lists of good reviews

Formulate hypothesis

- For ideas that still persist (very few do), now is the time to actually formulate your hypothesis. Be specific:
  - What is it you want to study
  - What is the exposure?
  - What is the outcome?
  - In whom?
  - Where?
  - Etc.
  - *Note: This is not a semantic exercise, write them however you want*

Feasibility

- Naturally, feasibility will have been at the back of your head throughout all of the planning, but I don't think any study concept should fall on feasibility until you've done the following:
  - Thought about the relevance of the study
  - Done the background check
  - Formulated specific hypotheses
- Once you have these details clear, however, feasibility should be assessed before proceeding

Feasibility

- So, what should you consider when assessing the feasibility?
  - Can you get your hands on the data?
  - Will you have enough data (power)?
  - Will you be able to answer the question meaningfully with the available data?
  - Will the limitations of your envisioned design be forbidding?
  - Will you be able to get enough money?*
  - Will you have the time?*
  - Will the study be ethically feasible?
  - Most importantly, however, do you WANT to do the study?

Designing a study

- So, you've come up with an idea, realized that its worthwhile and that it is doable, great! Now, how do I actually do it?
- Often, the study design is obvious given the exposure/outcome combination or given the available data
- If not, consider the natural roles of the major study designs:
  - Rare disease – case-control study
  - Rare exposure – cohort study
  - Considerable indication bias – RCT

Designing a study

- In choosing a study design, keep the integral design limitations in mind:
  - Observational studies: confounding by indication!
  - Prospective cohort studies: time consuming…
  - Retrospective cohort studies: survival bias, lack of sufficiently detailed data
  - Case-control studies: selection bias* and recall bias
  - RCTs: time, hard work and expensive
  - Cross-sectional studies: temporal sequence

Prospective vs. Retrospective

- In choosing whether your studies will be prospective or retrospective, keep in mind:
  - While you will be able to tailor the data collection to your exact requirements, prospective studies are time consuming and expensive
  - At the same time, retrospective studies, which are much cheaper, depend on available data, making lacking data on covariates a considerable problem, and are also susceptible to selection bias

Fundamentals of study design

- Recall the counterfactual ideal:
  - We want to compare disease risk in individuals who would have been EXACTLY identical had they not received different exposures*
  - In other words, we want to avoid comparing individuals that are different in unmeasureable ways
- This is implicit in RCTs

Deciding your study population

- Often, the choice of study population is obvious, but there are some pointers:
  - Maximize internal validity: find a homogenous population with little variation in important confounders
  - Maximize external validity: make sure there is sufficient variation in the exposure of interest
  - Maximize power: make sure the outcome of interest is sufficiently common

IS THIS POSSIBLE?

Deciding your study population

- Keep in mind that you want to have a study population that is motivated for enrollment (and persistence)
- You also need do be able to follow your study population, or contact them again for more information if needed
- Finally, how will you find out whether the outcome of interest occurred?

Getting the data

- How will you get your hands on the data?
  - Exposure ascertainment
  - Outcome ascertainment
  - Picking the covariates (and measuring them)
    - The choice of confounders and possible effect modifiers to include is sometimes difficult, but requires COMPLETE attention
    - It is usually better to ask about more than you think (but maybe not too much)*
    - For every tentative covariate, ask yourself: What will I do with this piece of information?

Exposure ascertainment

- How will you collect exposure data?
  - Interviews – who will do them?
  - Medical records – do they contain enough data?
  - Questionnaires – will the patients be able to complete them?
  - Registers – are they complete enough?
  - Other?
- Generally, the exposure ascertainment is similar in case-control and cohort studies

Exposure ascertainment

- Key differences between cohort and case-control:
  - Case-control: Is it a rapidly fatal (or invariably fatal) disease? Rapid exposure ascertainment!
  - Case-control: Is there opportunity for recall bias? Manage it!
  - Cohort study: If there is opportunity for exposure to change? Measure exposure repeatedly!
  - Cohort study: Will you be able to capture loss to follow-up?

Outcome ascertainment

- The ascertainment of outcome (i.e. disease occurrence) is a crucial aspect of conducting an epi study.* You need to figure out:
  - How and where will you find your cases?
  - Is it possible to know wherefrom these cases arose?
  - Will the case ascertainment be rapid and complete enough for your purposes?
  - Will you be allowed to study these persons?

Summarizing your study plan

- Once you have a clear idea of what it is you want to do and how, its time to write your Study protocol!
- The study protocol is your (detailed) plan for conducting the study.
- It should outline the rules for the conduct of your study to avoid variations*
- In some ways, it is similar to a grant application, but usually there is more detail (and less selling) in the protocol

Study protocol

- Typically, protocols follow a clear structure:
  - Background and rationale
  - Specific aims
  - Study design
  - Study population
  - Exposure and outcome ascertainment
  - Power analysis*
  - Statistical analysis*
  - Human subjects and ethical considerations
  - Time plans

Manual/standard operating procedures

- In addition to the study protocol, it is customary (at least in larger studies) to keep a batch of SOPs that ensure consistency in:
  - Data collection (i.e. interviewing, medical record review, etc.)
  - Data entering
  - Data management
- The goal is to minimize bias by design and misclassification

Hints for data collection

- Get more data than you think you need
- When designing questionnaires and interview sheets, take your time… plenty of time… much more time than you think… and test them!
- In all measurements, avoid categorizing the questions beforehand (if you can)
- For binary variables like death yes/no, AMI yes/no: don't forget the exact date!
- Always make a distinction between no/unknown/missing
- In interviews, questionnaires and record review, always attempt to blind the interviewer/data enterer, reviewer

Summary

- The practicalities of realizing a study concept go through many steps with a fairly logical sequence
- While the feasibility and potential gain of a certain project is important, I would argue that the relevance of the study idea and your interest is most important
- In the initial stages of planning of a study, do not be too pessimistic