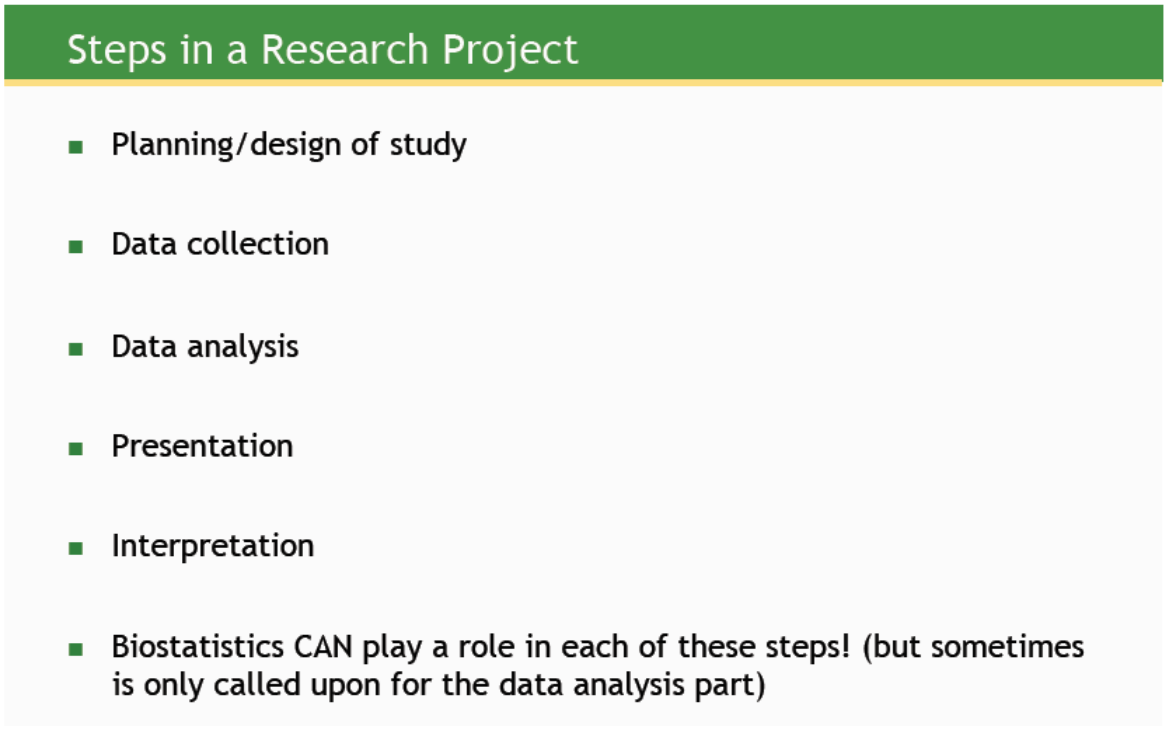
Transcript: Role of Biostatistics in the Steps of a Research Project

[Acknowledgement]



What are the steps involved in a typical research project in public health, medicine, or any other discipline?

What roles can biostatistics play in each step?

Typically from start to finish the steps include

1) Planning or designing the study

2) Data collection where we gather, record, and check the needed information

3) Data analysis

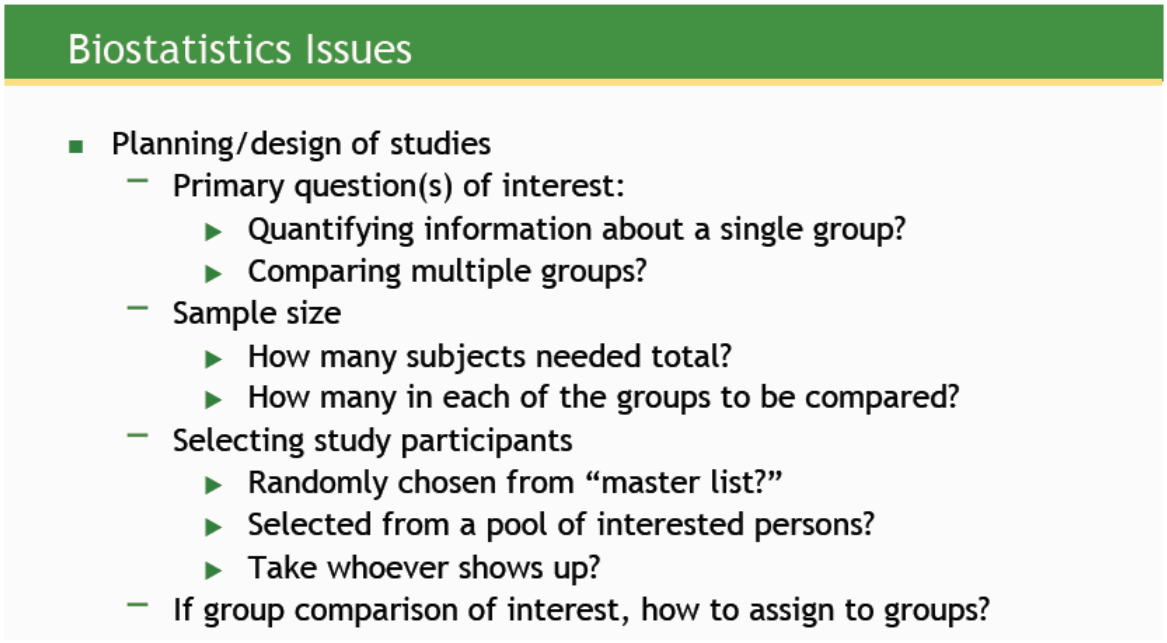
4) Presentation of results

and

5) Interpretation of the results in context

Not too surprisingly, biostatistics can play a role in each of these steps.

But sometimes unfortunately it's only called upon for the data analysis part when it's actually too late to rectify any mistakes made earlier on.



First, Step 1)

In the planning and design phase of a study, biostatistics, biostatisticians, and those trained in biostatistics can help with identifying the primary research questions of interest and turning these into measureable variables.

How can we quantify information about a single group?

Do we want to compare multiple groups?

How can we quantify the things we want to compare?

What can be quantified?

What can't be quantified but is still important to record?

Biostatistics can help with determining the sample size.

How many subjects do we need all together in this study?

If we're comparing groups how many do we need in each group?

Biostatistics could help with choosing the study participants.

Are we lucky enough to be able to randomly choose them from some master list
of everyone in the population we want to study?

Are we going to randomly select them from a pool of interested persons who agreed to be in the study if selected?

Are we going to take whoever shows up?

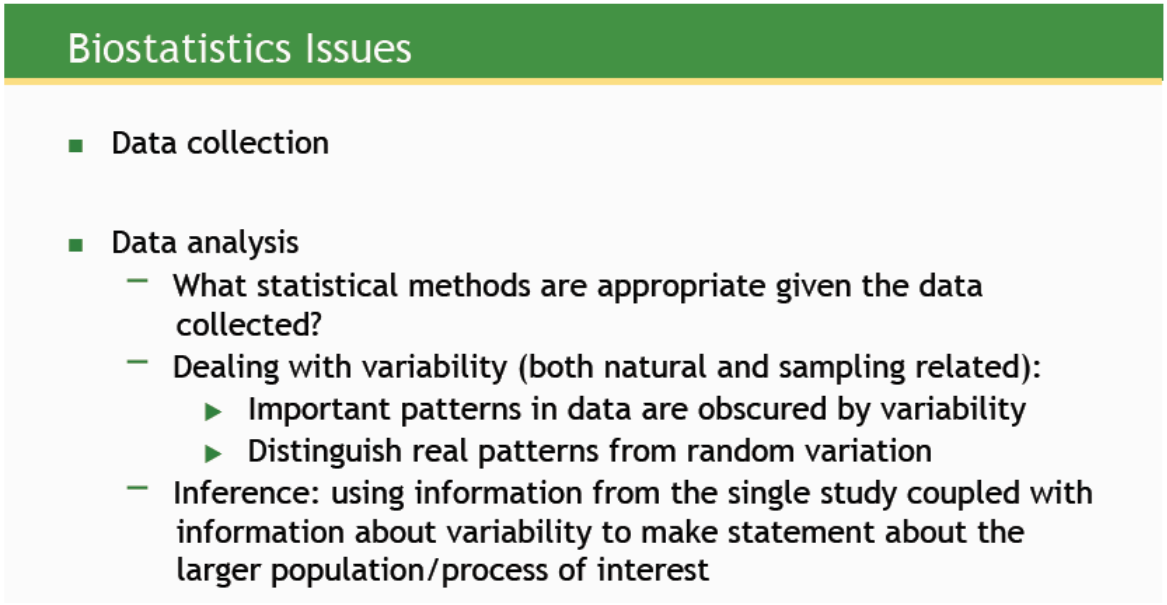
All of this depends of course on the context of what we're doing, how easy it is to get study subjects, and a lot of other concerns.

But we want to think about these kinds of issues early on because it may have implications for how we conduct the study and analyze the data.

And if a group comparison is of interest, for example, suppose we want to compare the benefits of two preventative treatments for the flu.

How are we going to assign participants to one of the two groups?

Biostatistics can play a role in that.



Step 2) Biostatistics can also play a role in the data collection.

Unfortunately, frequently biostatisticians are not involved in this aspect but even if we cannot bring anything new to the table in terms of going out and collecting the data itself, it is still very good for those who are going to be analyzing the data and interpreting the results to have some sense of how it was collected.

What difficulties there might have been in collecting certain pieces of information and how things could be better measured in a subsequent study.

So, it's a good idea for statisticians to be involved in the data collection process to inform the other steps, although frequently we are not, either by choice of ourselves or the primary investigator.

Step 3) Certainly biostatistics has many roles in the data analysis portion.

We want to determine what statistical methods are appropriate given the data we've collected.

We want to determine how to deal with the variability in our data - both natural variability - which we'll discuss and sampling related variability - which is a function of the natural variability.

We will be discussing these ideas throughout this course.

Many times what we're trying to do is distinguish important patterns in our data that may be obscured by some variation in the data.

In essence, we're trying to separate the signal in our data from the noise around it.

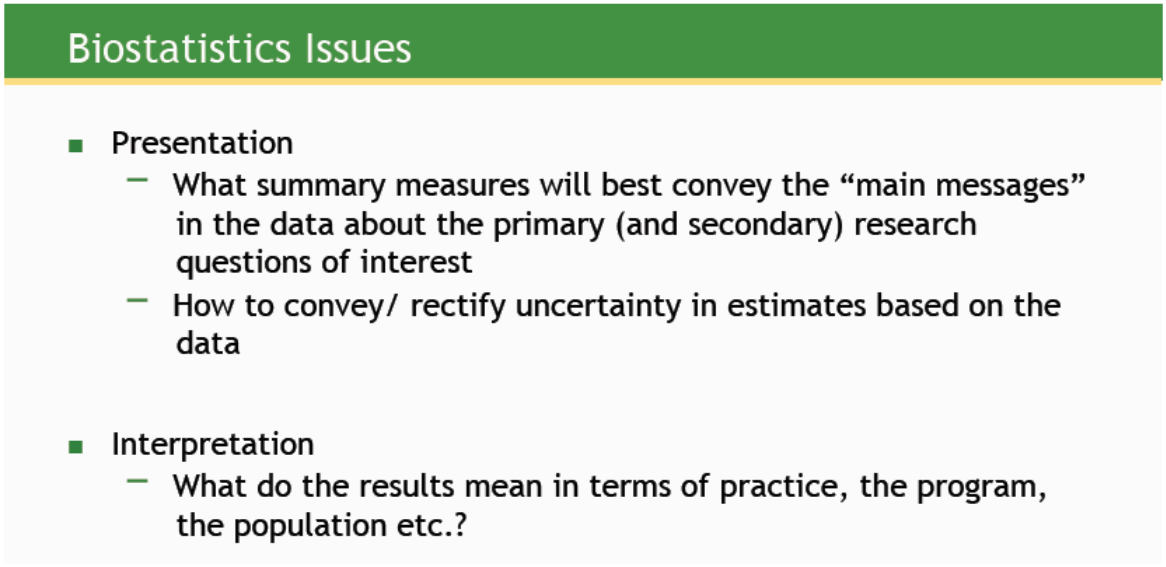
We're trying to distinguish real patterns, real differences, real findings from those that are just random noise.

We'll be talking a lot about that this semester.

And in that same spirit we want to perform something called inference.

Using information from a single study on an imperfect subset of subjects to answer questions about the larger population from which the subset was taken.

We use the information from our single study combined with information about variability to make inferences about the larger population.



Step 4)

Biostatistics can help in the presentation.

Determining which summary measures will best convey the main messages in the data.

How to best present the research results about the primary and secondary research questions of interest.

And how to convey or rectify uncertainty in estimates based on the data.

And finally, Step 5)

This is one of the most important facets of research, interpreting results.

Certainly this is a job for everyone involved in the research not just those dealing with the statistical aspects but it requires a knowledge of statistics even if you're not the primary statistician.

What do the results mean in practice, the program that you are evaluating, the research that you're doing, the population you're studying, the policy issues etc.

So interpretation is a big part of this and relies on having some understanding of the methods used and measures presented.

Steps in a Research Project

- Planning/design of study
- Data collection
- Data analysis
- Presentation
- Interpretation
- Biostatistics CAN play a role in each of these steps! (but sometimes is only called upon for the data analysis part)

So hopefully now you see that biostatistics can play a full role in any type of research undertaking you may embark on or in understanding the results of research studies related to your area of interest.