Stats 101 for health

29th October 2020 Ballarat Health Services.

This presentation is an overview intended to provide an introduction to concepts. If you have technical questions please contact BIRCH (admin@birch.org.au)

Neroli Sawyer
BA Psyc, Applied Statistics (Hons Statistics), PhD (Epidemiology)

What usually happens

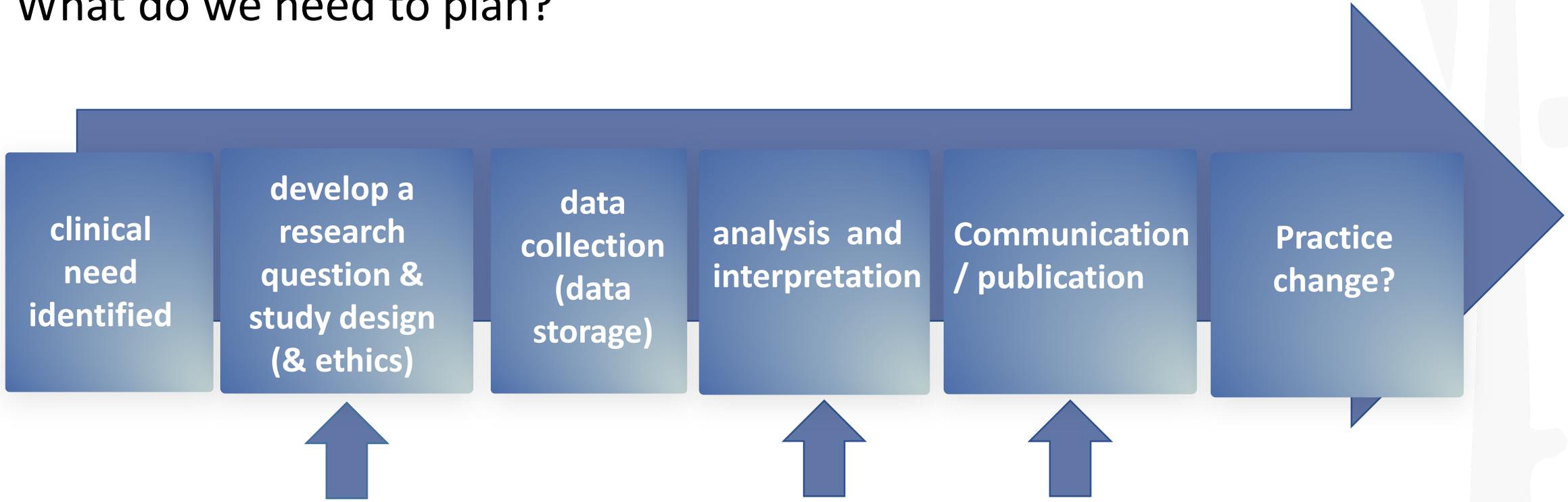| Clinician has a research question		Collects data		Seeks statistics support help analyse data
	→		→	

Risks: question may not be answered
- if the correct data has not been collected
- not enough data
- incomplete data
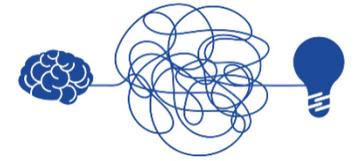
Lets have a look at the research Journey

What do we need to plan?

| clinical need identified | develop a research question & study design (& ethics) | data collection (data storage) | analysis and interpretation | Communication / publication | Practice change? |

This is where a statistician can help

Who is BIRCH currently collaborating with?

- Clinicians who have observed issues and are investigating patterns and relationships that could inform care (e.g. between injury-diagnosis, disease-treatment)

- They are working on and developing research projects based on the data they can access

- Each investigation, research question will require different types of data

The type of analysis you can use to answer your research question depends on what form your data takes.

Know your data: the DV and IV

DV

Dependent variable

The variable being tested and measured in an experiment, and is 'dependent' on the independent variable e.g. how does a patient's *temperature* change in response to different forms of medication

IV

Independent variable

The variable the experimenter changes or controls and is assumed to have a direct effect on the dependent variable e.g. if we *change the medications* what effect does that have on a patient's temperature

Know your data: 4 scales of measurement

Categorical

qualitative data

1. Nominal scale – How the experimental unit (e.g. a patient) is categorised e.g. gender (male/female) or ward (medical/surgical).

has no natural order

2. Ordinal scale – categorisation that has a natural order e.g. level of risk (low, med, high).

Numerical

quantitative data

3. Interval scale –there is order and the difference between two values is meaningful e.g. temp or pH levels. Has no true zero (does not imply absence e.g. of temp etc).

4. Ratio scale –there is a true zero. Unlike on the interval scale, a zero on a ratio scale means there is total absence e.g. Length, area, and population.

Two broad types of data analysis

Descriptive

The data is 'stand alone'

1. For **count data** (nominal or ordinal scales): report frequency & percentages within groups e.g. what is the frequency or number (n) and proportion or percentage (%) of females in your data set

2. For **discrete or continuous data** (interval or ratio scales): report means or medians for measures of centre; standard deviation, IQR and range for measures of variation (spread)

Inferential

The data is *representative* of a larger population

Enables generalisation of results to a larger population of interest

But only if certain criteria or assumptions are met (e.g. n > 25 – 30 for each group/cell)

NB: assumptions (not discussing further in this talk) include, but are not limited to, normality, equal variance, linearity

What type of analysis: descriptive or inferential?

This will depend on the purpose of your investigation

i.e. descriptive statistics generated from a descriptive analysis may be adequate for an internal audit or report

It will depend on the representativeness of your data

i.e. if you want to generalize your results or make predictions about a larger population of interest your data must be representative of that population

It will also depend on how much control you have over your sample size

i.e. even if you have a representative sample, inferential analyses may not be appropriate if your sample size is too small – you will not have enough power or meet basic assumption testing

Some examples of inferential analysis:

The type of inferential analysis you can do will depend on the nature of your research question, your study design and what data you have access to, or have collected.

Examples of study design:

1. Between or within groups tests (relationship between *1 numeric DV and 1 or more categorical IVs*) e.g. includes t-tests and ANOVAs

2. Chi-square tests (relationship between *categorical variables*) e.g. includes goodness of fit and independence tests.

3. Correlation and regression (relationship between *2 or more numeric variables*) e.g. includes simple and multiple regression

Study design: comparing 2 or more groups

DV = Numeric (e.g. hours or weight)

IV = Categorical (e.g. gender – males/females or time – before/after/6mth follow up)

Between groups e.g. is there a relationship between gender (male/female)and number of hrs/week watching Netflix? i.e. is there a difference between males and females in the average number of hrs/week they watch Netflix?

Within groups e.g. is there a relationship between weight change and a weight loss intervention (before/after/6mth)? i.e. is there a difference in weight before, after and at 6mths later, after a weight loss intervention/program?

DV: Dependent variable, IV: Independent variable

Study design: Chi-square test of independence

DV = Categorical (e.g. smoking– yes/no)

IV = Categorical (e.g. gender – male/female)

Is there a *relationship* between gender and smoking status?

OR, another way of phrasing this …

Are gender and choice to smoke *independent* of each other?

i.e. do males and females choose differently when it comes to smoking?

DV: Dependent variable, IV: Independent variable

Study design: Correlation/regression

DV = Numeric (e.g. number of falls/week)

IV = Numeric (e.g. cognitive score :MMSE)

Correlation e.g. tests if is there a positive or negative linear relationship between number of falls/week and MMSE?

Regression e.g. models the relationship (above) and then, is able to predict falls based on MMSE

DV: Dependent variable, IV: Independent variable

BIRCH

BIRCH research mentors based in Ballarat

- Biostatistics
- Honorary Assoc Prof Unimelb,
  Rural health MBBS, FRCGP
- Former Associate Dean of Research,
  Unimelb, Faculty Medicine,
  Dentistry & Health Sciences

Approachable, conversational, in-confidence