Key Considerations in Quantitative Research for Writing & Publication

Bishwajit Ghose

MPH (Nutrition), PhD (Epidemiology)
Postdoctoral fellow, School of Nursing, uOttawa

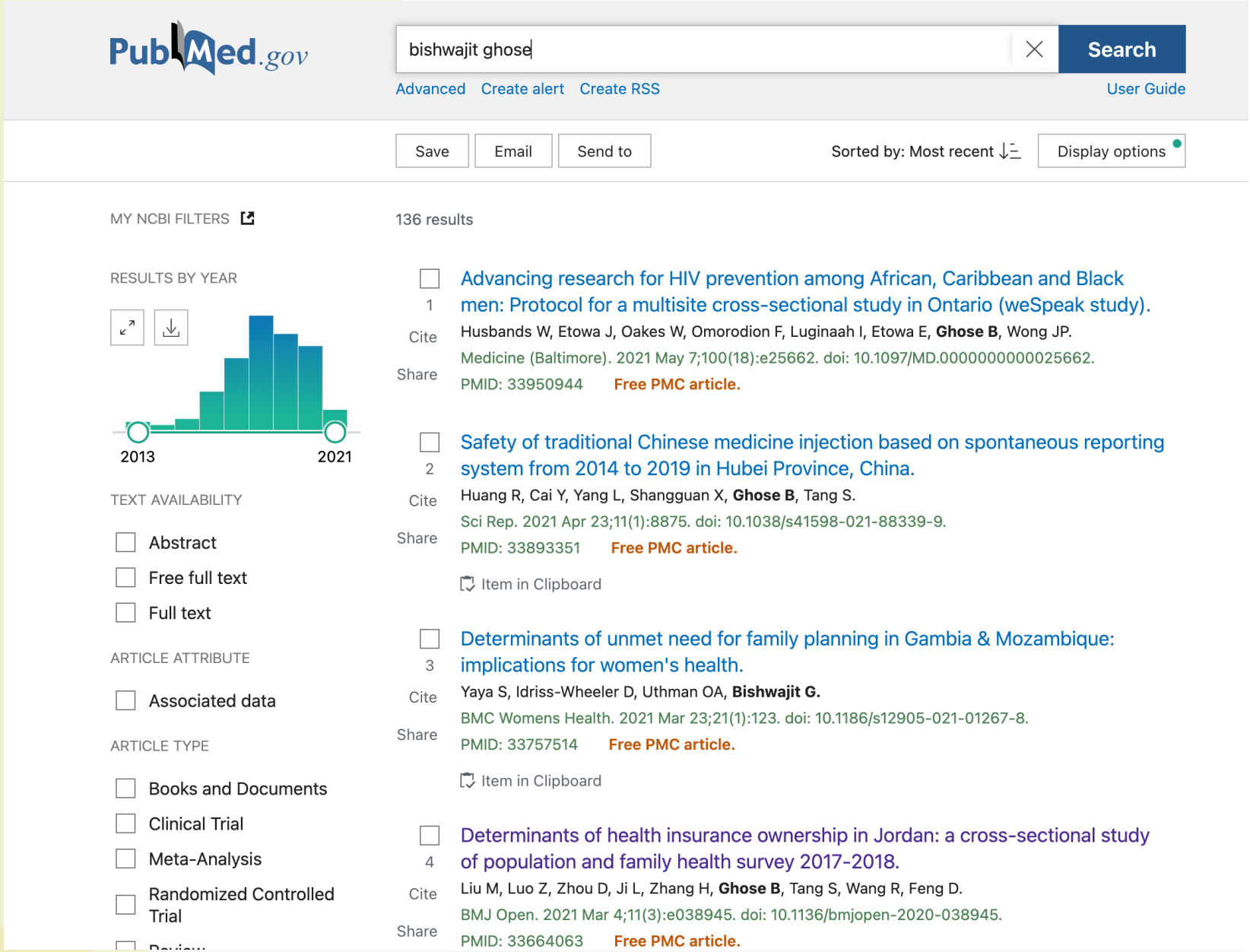
Academic editor

Women

PLOS ONE
Biomed Research International
BMC Pregnancy and Childbirth

Interests

Research Methodology
Social epidemiology
Big data
Machine learning



Learning objectives

Core concepts in Quantitative research

- Research Design
- Statistical Inference

Understanding data & Data analysis.

- Understanding Variables
- Choosing the Right Methods

Research & publication

- Life cycle of publication
- Types of research papers
- Standard structure of a manuscript

ELSEVIER

Journal of Medical Imaging and Radiation Sciences 45 (2014) 201-203

Journal of Medical Imaging and Radiation Sciences

Journal de l'imagerie médicale et des sciences de la radiation

www.elsevier.com/locate/jmir

Guest Editorial

Five Percent Is Not Enough! Why We Need More Qualitative Research in the Medical Radiation Sciences

Amanda Bolderston, BSc, MSc, RTT, FCAMRT*

British Columbia Cancer Agency, Fraser Valley Centre, Surrey, British Columbia, Canada

Research is one of the primary hallmarks of a profession. Medical radiation technologists (MRTs) are part of a "neophyte academic profession" [1] that has seen a relatively slow uptake of research and scholarly enquiry [2]. With the shift to baccalaureate education (and beyond) and the gradual increase in MRTs obtaining academic appointments, more practitioners are becoming research-active. Our knowledge base is still heavily reliant, however, on work done by other professions such as medicine, physics, and nursing [2]. There are important profession-specific gaps that need to be filled to inform practice, policy, and education which should be addressed by MRTs to establish a unique and distinct paradigm for our professional research.

When we do conduct research, we tend to focus on quantitative research that "reduce(s) experience to well-defined variables…for investigation" [3] or work that uses experimental designs, involves statistical analysis and deals with numbers as data. Our publications reflect this, for example only about 5% of submissions to the *Journal of Medical Imaging and Radiation Sciences* (*JMIRS*) in the last 8 years have been nonquanti-

As a further potential barrier, for a novice researcher, the approaches and methodologies of qualitative research can be a whole new language. Many MRTs have done statistics courses and may vaguely remember the difference between causation and correlation. How many of us are comfortable with the concepts and lexicon of qualitative research such as the differences between Husserlian and Heideggerian phenomenology? I am obviously exaggerating to make a point, but the roots of qualitative research lie in the traditions of the social sciences. Those of us in the applied health sciences, without a background in anthropology, sociology, or psychology have struggled to learn the rules and jargon of an approach that sometimes seems to be situated "inaccessibly high up in the misty mountains of academic discourse" [5].

It has not helped in the uptake of qualitative research that, until fairly recently, it has often been measured against quantitative research and been found wanting. The two approaches were traditionally posited as being diametrically opposed. In health care, the subjectivity and descriptive nature of the

OPEN ACCESS Freely available online

PLoS one

Is Qualitative Research Second Class Science? A Quantitative Longitudinal Examination of Qualitative Research in Medical Journals

Kerem Shuval[1]*, Karen Harker[1,2], Bahman Roudsari[1], Nora E. Groce[3], Britain Mills[1], Zoveen Siddiqi[1], Aviv Shachak[4]

1 School of Public Health, Division of Epidemiology, Human Genetics and Environmental Sciences, University of Texas at Dallas, Dallas, Texas, United States of America, **2** University of Texas Southwestern Medical Center, Dallas, Texas, United States of America, **3** The Leonard Cheshire Disability and Inclusive Development Centre, Department of Epidemiology and Public Health, University College London, London, United Kingdom, **4** Faculty of Information and Department of Health Policy Management and Evaluation, University of Toronto, Toronto, Canada

Abstract

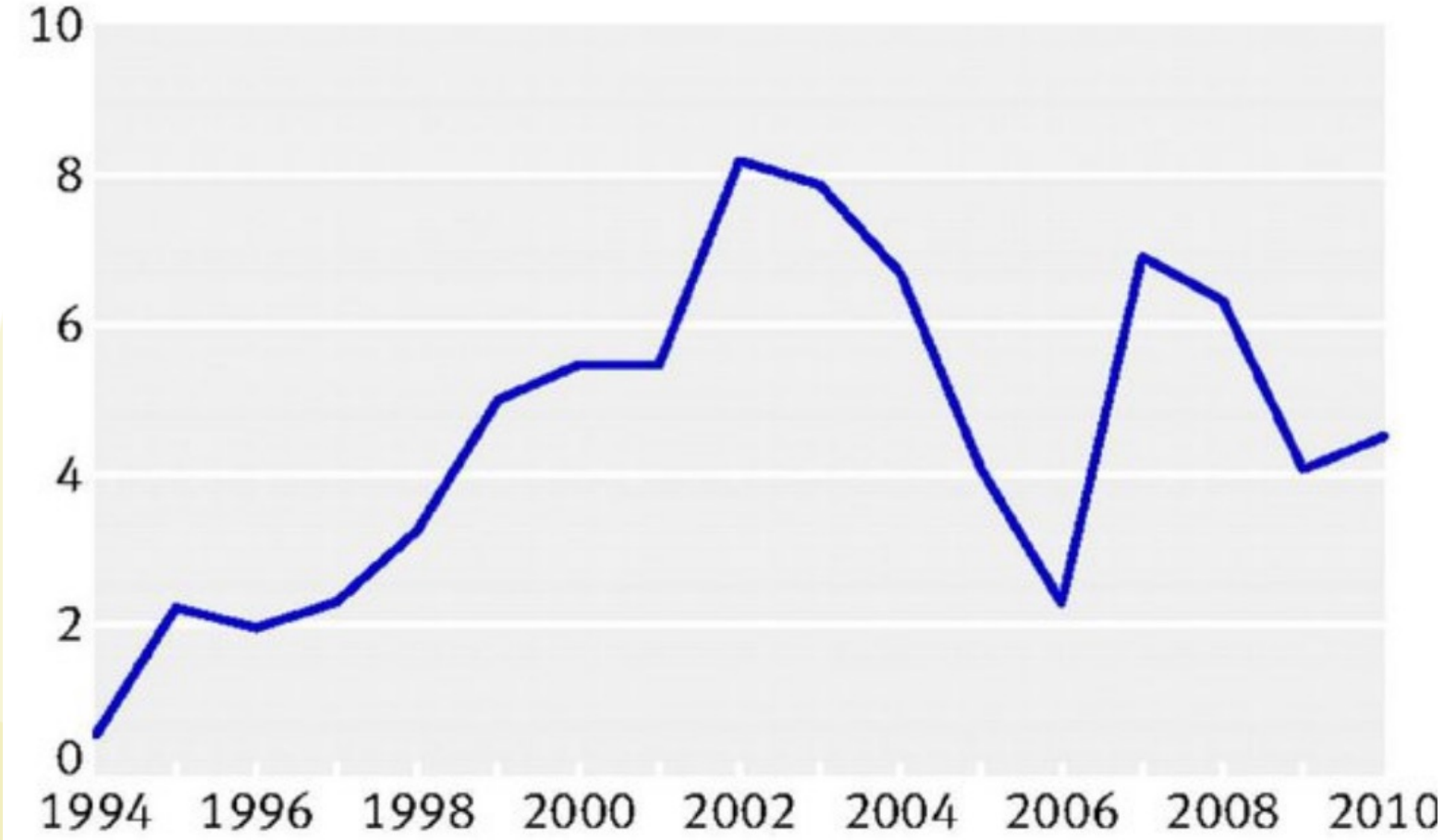
Background: Qualitative research appears to be gaining acceptability in medical journals. Yet, little is actually known about the proportion of qualitative research and factors affecting its publication. This study describes the proportion of qualitative research over a 10 year period and correlates associated with its publication.

Design: A quantitative longitudinal examination of the proportion of original qualitative research in 67 journals of general medicine during a 10 year period (1998–2007). The proportion of qualitative research was determined by dividing original qualitative studies published (numerator) by all original research articles published (denominator). We used a generalized estimating equations approach to assess the longitudinal association between the proportion of qualitative studies and independent variables (i.e. journals' country of publication and impact factor; editorial/methodological papers discussing qualitative research; and specific journal guidelines pertaining to qualitative research).

Findings: A 2.9% absolute increase and 3.4-fold relative increase in qualitative research publications occurred over a 10 year period (1.2% in 1998 vs. 4.1% in 2007). The proportion of original qualitative research was independently and significantly associated with the publication of editorial/methodological papers in the journal ($b = 3.688$, $P = 0.012$); and with qualitative research specifically mentioned in guidelines for authors ($b = 6.847$, $P < 0.001$). Additionally, a higher proportion of qualitative research was associated only with journals published in the UK in comparison to other countries, yet with borderline statistical significance ($b = 1.776$, $P = 0.075$). The journals' impact factor was not associated with the publication of qualitative research.

Conclusions: Despite an increase in the proportion of qualitative research in medical journals over a 10 year period, the proportion remains low. Journals' policies pertaining to qualitative research, as expressed by the appearance of specific guidelines and editorials/methodological papers on the subject, are independently associated with the publication of original qualitative research; irrespective of the journals' impact factor.

% of Qualitative Research Published by BMJ



Quantitative research

- …is a means for testing objective theories by examining the relationship among variables *(Polit and Hungler 2013).*
- …is one in which quantitative data on **variables** are collected and analyzed.
- …is able to identify the deep nature of realities, its system of **relations**, its dynamic structure.
- …can determine the strength of association or correlation between variables, the generalization and objectification of the results through a sample that makes inference to a population.
- … can make causal inferences which explains why things happen or not in a particular way.

++ sides of Qn Research

There is always a better way to do the same thing

Place for creativity

Use of technology

Research design

… refers to the overall strategy to integrate the different components of the study in a coherent and logical way to answer the research question.

…constitutes the blueprint for the collection, measurement, and analysis of data.

Elements:

- Purpose statement
- Techniques for data collection
- Methods for research analysis
- Type of research methodology
- Probable objections to conducting research
- Research study settings
- Timeline
- Analysis measurement

Research design

Experimental

Aim:
To establish a cause-and-effect relationship an independent and dependent variable.

Type: Prospective e.g. individuals are followed time and data about them is collected as the outcomes change.

Example:
We want to test the effectiveness of dementia Patients is randomly divided into three groups: 1) receiving a high dosage of the drug (experimental group), 2) receiving a low dosage (experimental group), 3) receives a placebo (control group). After administering the drug for a period of time, if the condition of the experimental group subjects improved significantly more than the control group subjects, we can say that the drug is effective.

Quasi- Experimental

Aim:
To establish a cause-and-effect relationship between an independent and dependent variable.
N.B. randomization not possible due to ethical or practical reasons.

Example:
We want to test the effectiveness of a new care model e.g. empathetic model which is hypothesized to lead to shorter hospital stays among dementia patients. You choose two similar groups of physicians/nurses in two different hospitals one of which implements the empathetic model while the other does not. By comparing the length of hospital stays between the two hospitals we can find out whether the empathetic model is actually effective in reducing the length of hospital stay.

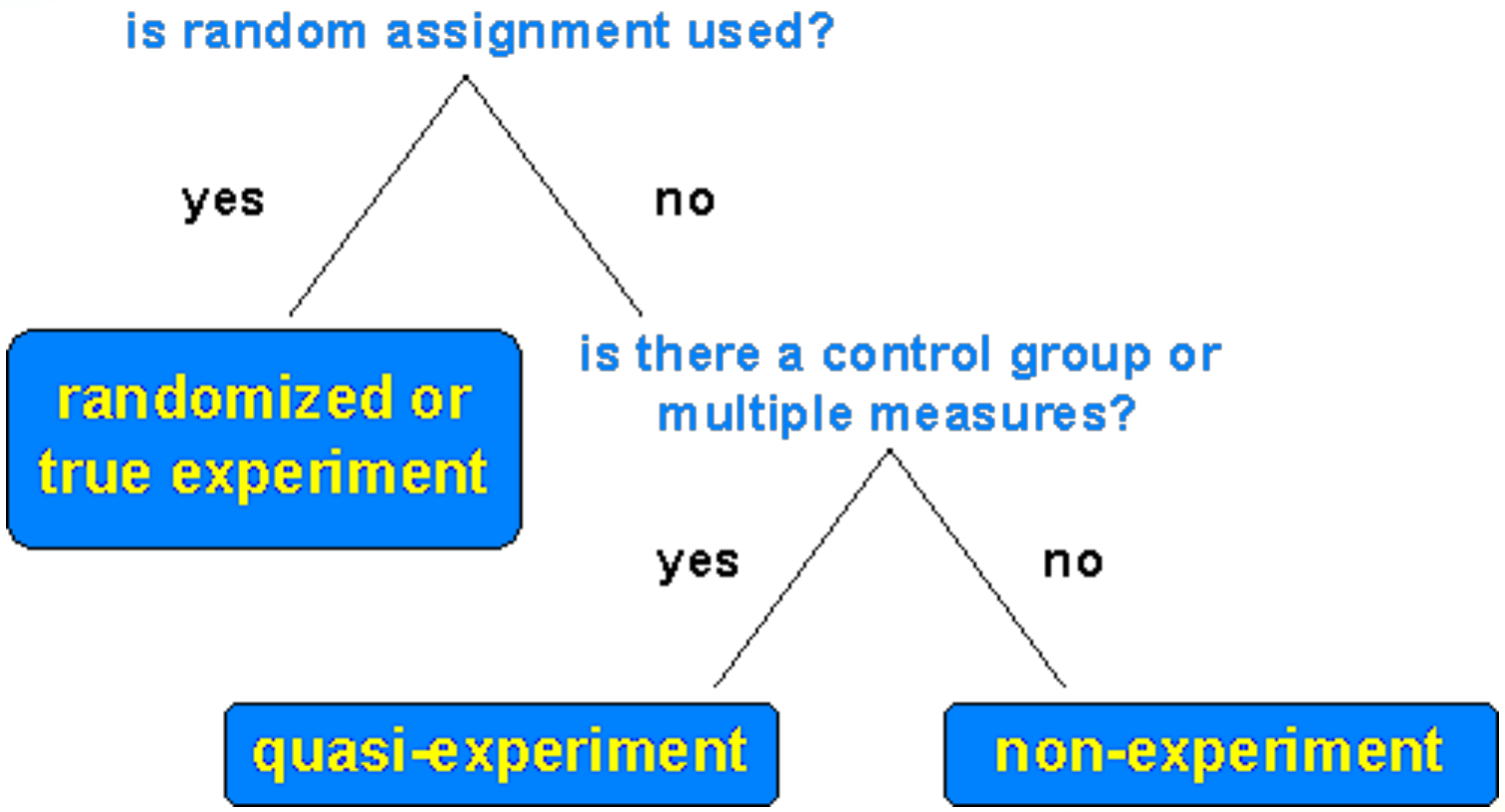
Observational

Aim: To demonstrate temporal association/ correlation between an independent and dependent variable.

Type: retrospective e.g. starts with an outcome then traces back to investigate exposures.
Requirements: Cases and controls must match by e.g. age, sex, recruitment methods.

Example: We want to assess whether lower smoking (the exposure) is associated with a higher risk of lung cancer (the outcome). Data are collected from hospital patients wo are smokers, and with and without lung cancer to compare the associations.

It is a matter of CONTROL…



Statistical inference

Statistical inference is the process of making inferences about a population based on certain statistics calculated from a sample of data drawn from that population.

1. *Point Estimation*

   Estimates an unknown parameter using a single number e.g. % of diabetic population.

2. *Interval Estimation*

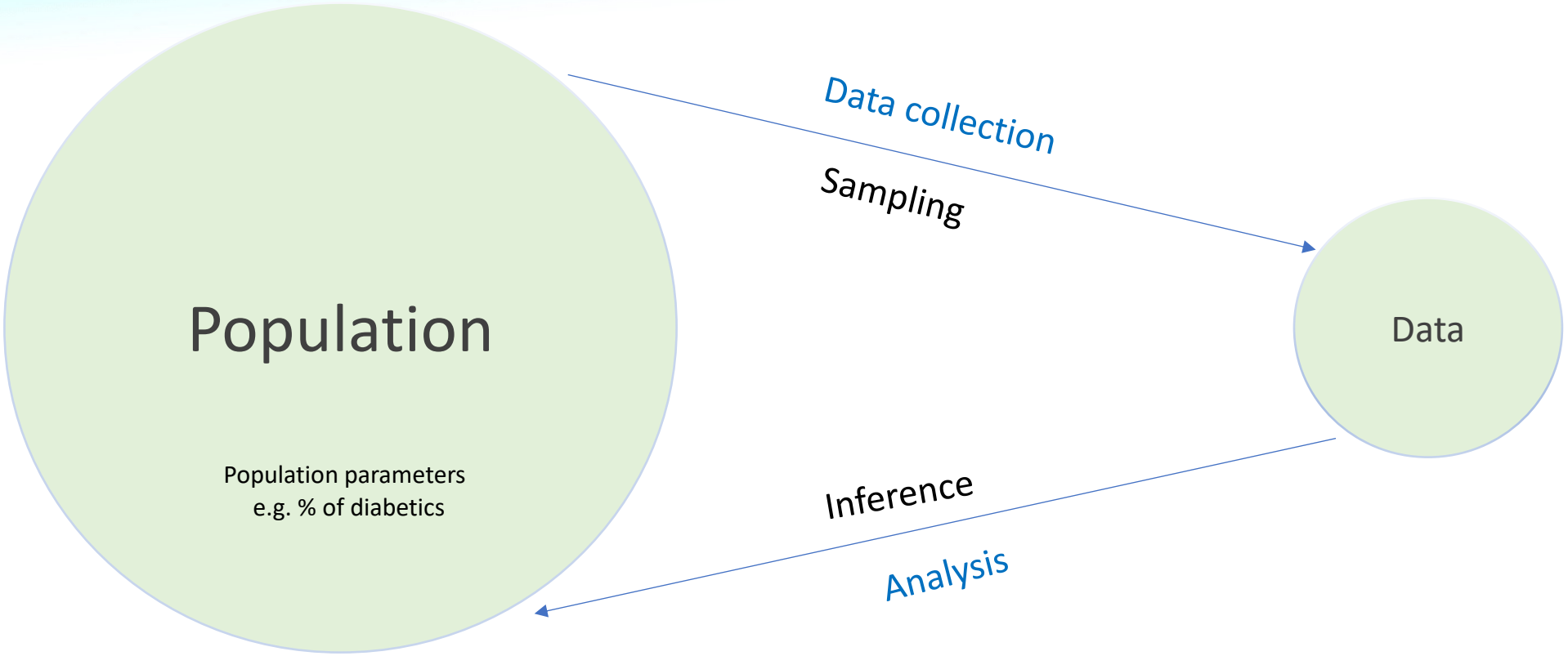
   Estimates an unknown parameter using an interval of values that is likely to contain the true value of that parameter e.g. % of diabetic population is between 4.2% and 5.7%.

3. *Hypothesis Testing*

   In hypothesis testing, we begin with a claim about the population (we will call the null hypothesis), and we check whether or not the data obtained from the sample provide evidence AGAINST this claim.

Statistical inference

Two steps:
1) Data collection & 2) estimation



Specify hypothesis

Primary question/hypothesis:
Empathetic nursing can improve health outcomes.

Secondary question/hypothesis:
If yes, in what population group?
e.g. Disease category, Immigrants.

Generate data

Primary surveys:
Face to face/telephone/ online survey.

Electronic health records:
e.g. iPatientCare.

Calculate statistics

Data cleaning:
e.g. outliers, missing values.

Data analysis:
e.g. descriptive (e.g. mean, prevalence), inferential (e.g. regression) methods.

Goodness of fit statistics, reliability, validity, sensitivity, specificity.

Visualize

Pie chart:
To show how much each category represents as a proportion of the whole by using a circular format.

Bar chart:
To show absolute values/proportions for each of the categories by using rectangular bars.

Understanding variables

- Variables are like **currencies** in quantitative analysis. Know them well!
- There are a few. Choice of statistical tests is dependent on the type of variables.
- Statistical methods are almost always specific to data types. Thus, choosing an inappropriate model will not work, or lead to flawed estimations.

Two things to know about variables:

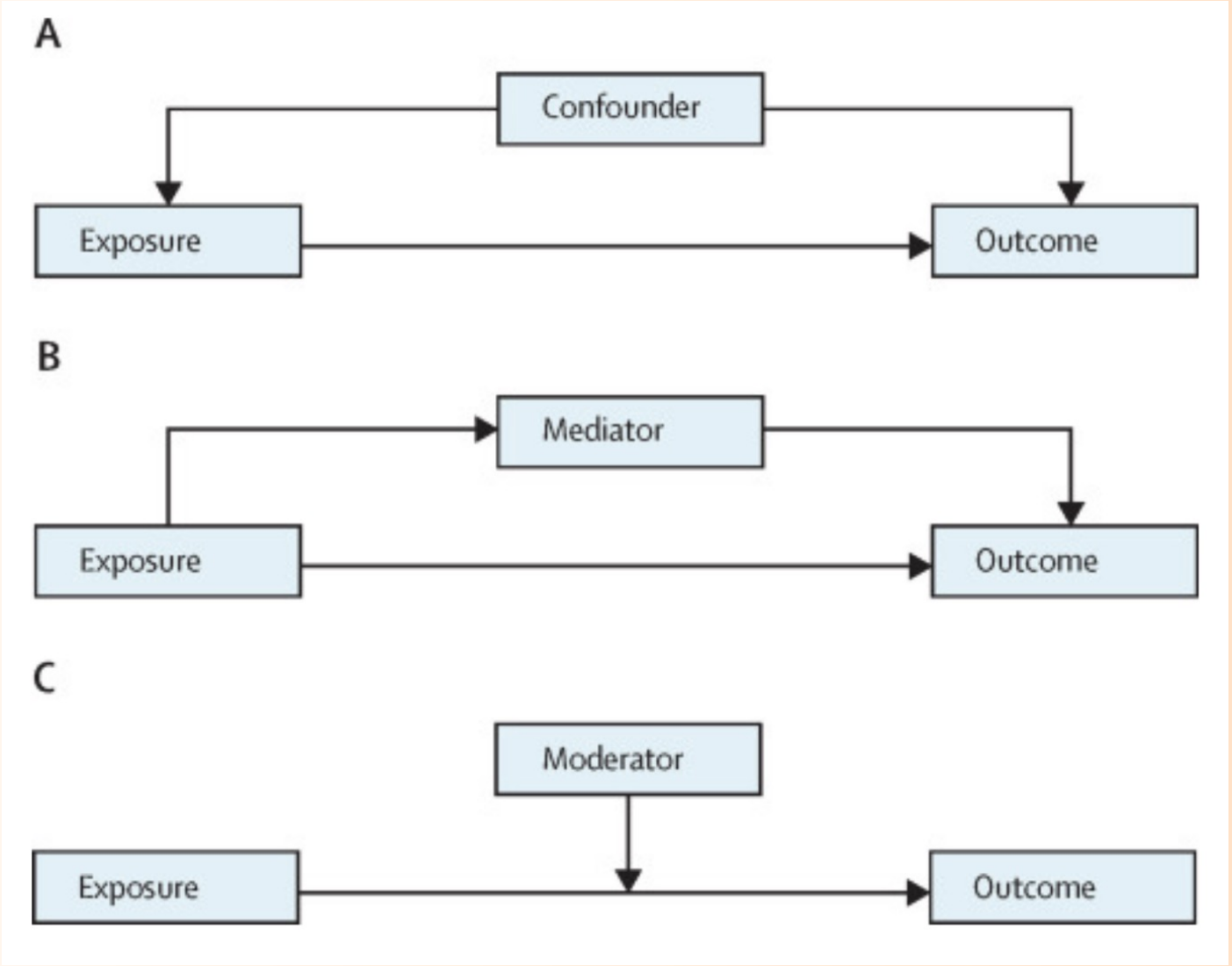
What role they assume in the equation:

e.g . **dependent (DV), independent** (IV), control, confounding, mediating, moderating.

What they look like:

e.g. numeric, categorical,

Role a variable may assume in the equation:



Dependent, independent & control

Does red light affect plant growth?

Independent Variable
Light Color

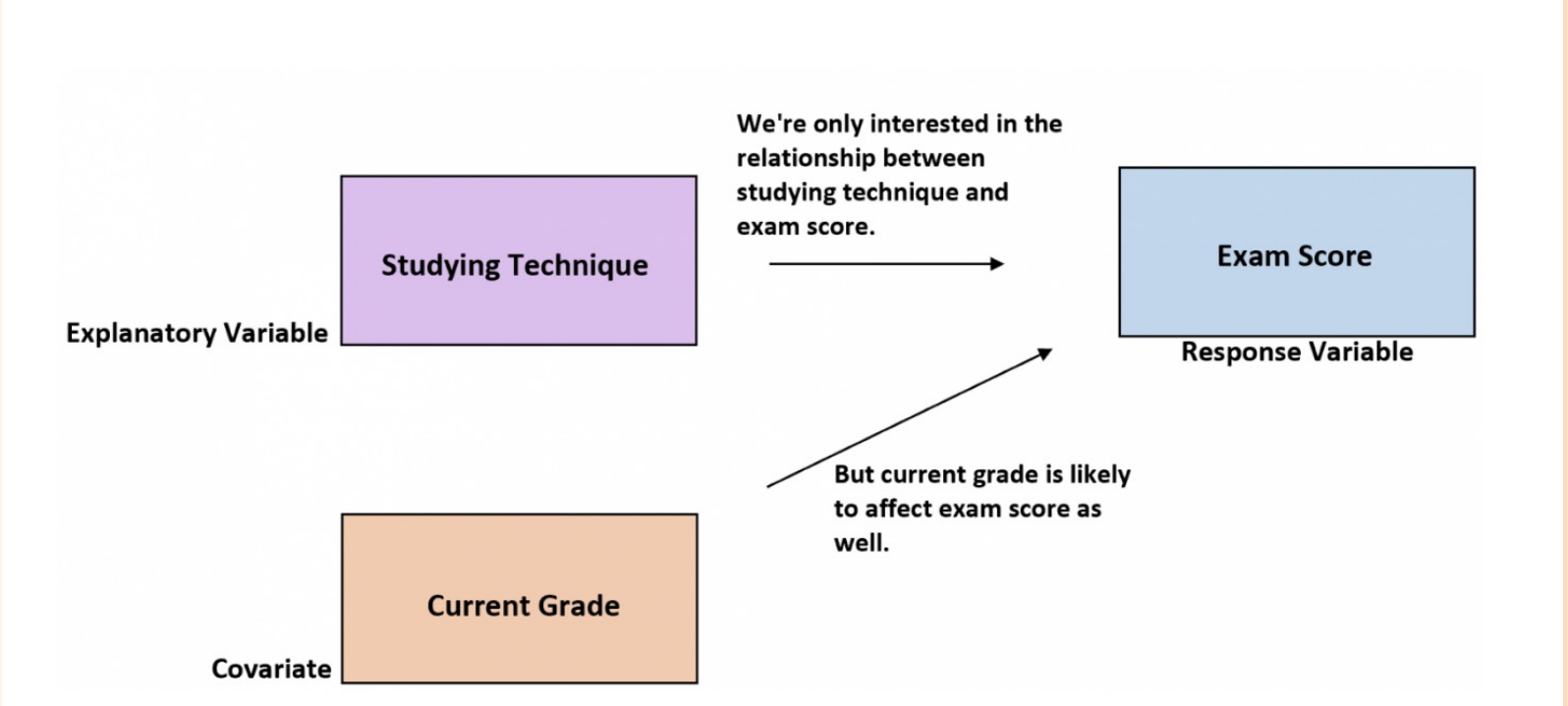
Dependent Variable
Plant Height

Control Variables
Water
Soil
Temperature
Day/Night Time

Covariate

Usually unwanted (not of primary interest, but better to be added).
Adding a covariate to a model can increase the accuracy of the estimations.

E.g. Current grade is actually not a variable of interest in this study; nonetheless, it can be included as a covariate so that researchers can see if studying technique affects exam scores even after accounting for the student's current grade in the class.

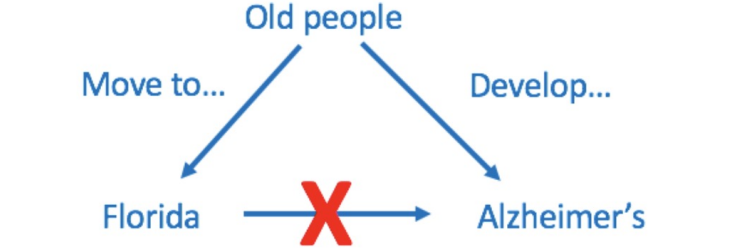


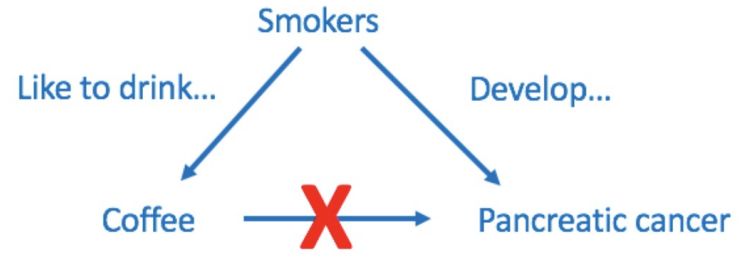
Confounding factor

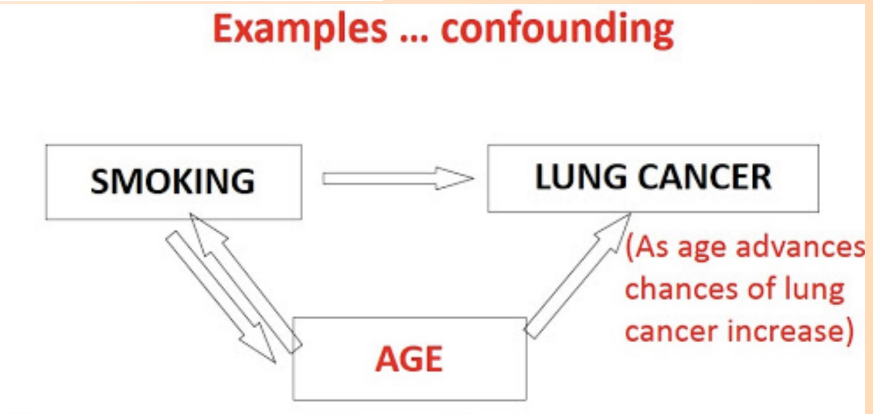
An independent variable represents the suppose cause, while the dependent variable is the supposed effect.

A confounding variable is a third variable that influences both the independent and dependent variables. Failing to account for confounding variables can cause you to wrongly estimate the relationship between your independent and dependent variables.

A confounder is a variable that influences both the dependent variable and independent variable, causing a spurious association.



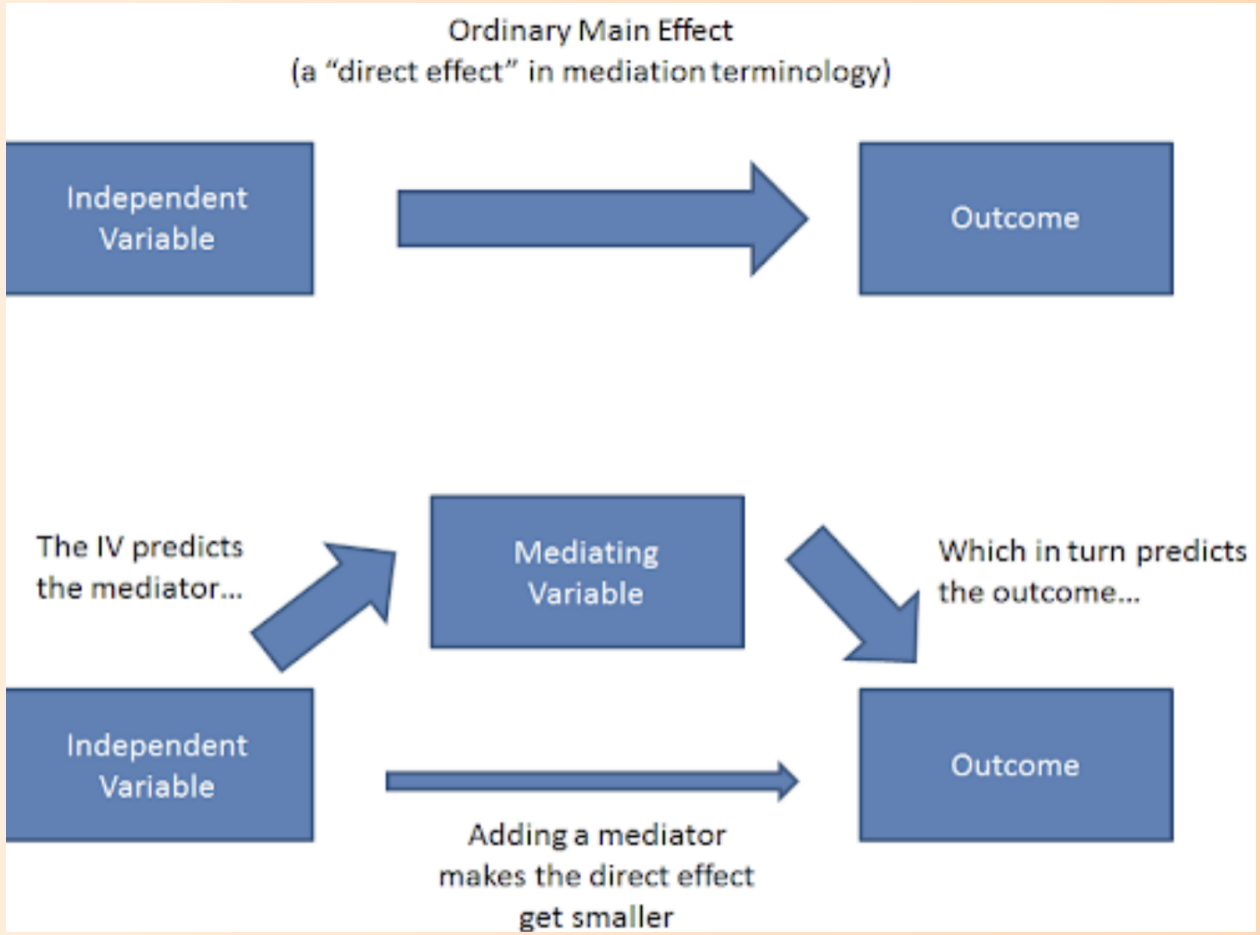




Mediating variable

- A mediating variable explains the relation between the independent and the dependent variable. It explains how or why there is a relation between two variables. A mediator can be a potential mechanism by which an independent variable can produce changes on a dependent variable. When you fully account for the effect of the mediator, the relation between independent and dependent variables may go away.

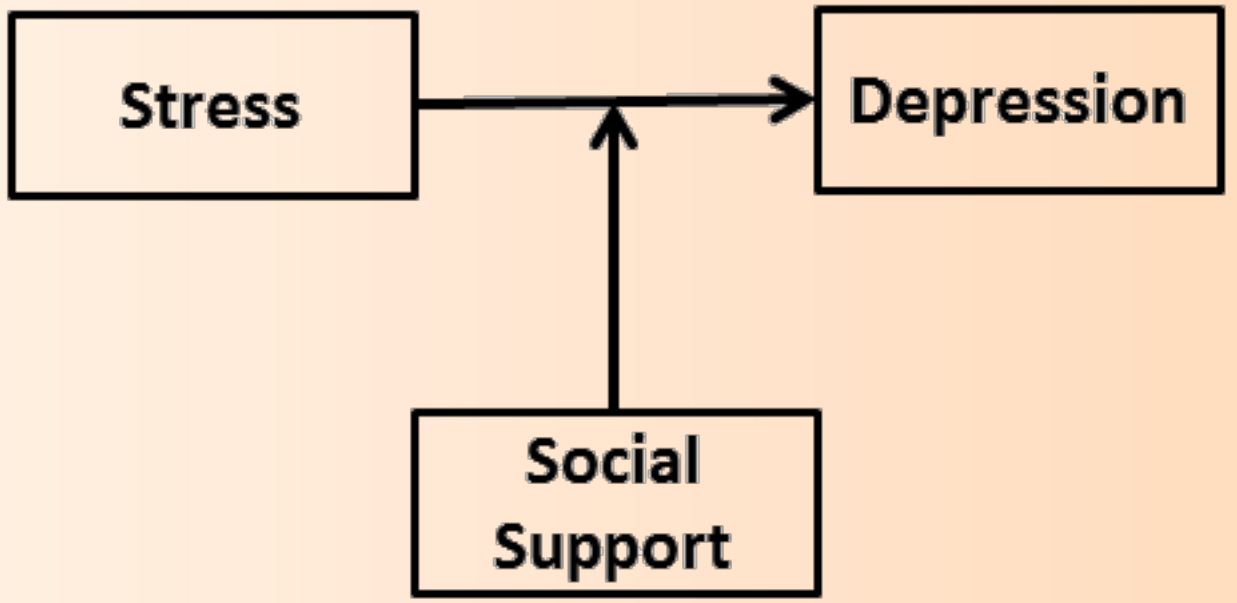
E.g. *you find a positive association between note-taking and performance on an exam. This association may be explained by number of hours studying, which would be the mediating variable.*



Moderating variable

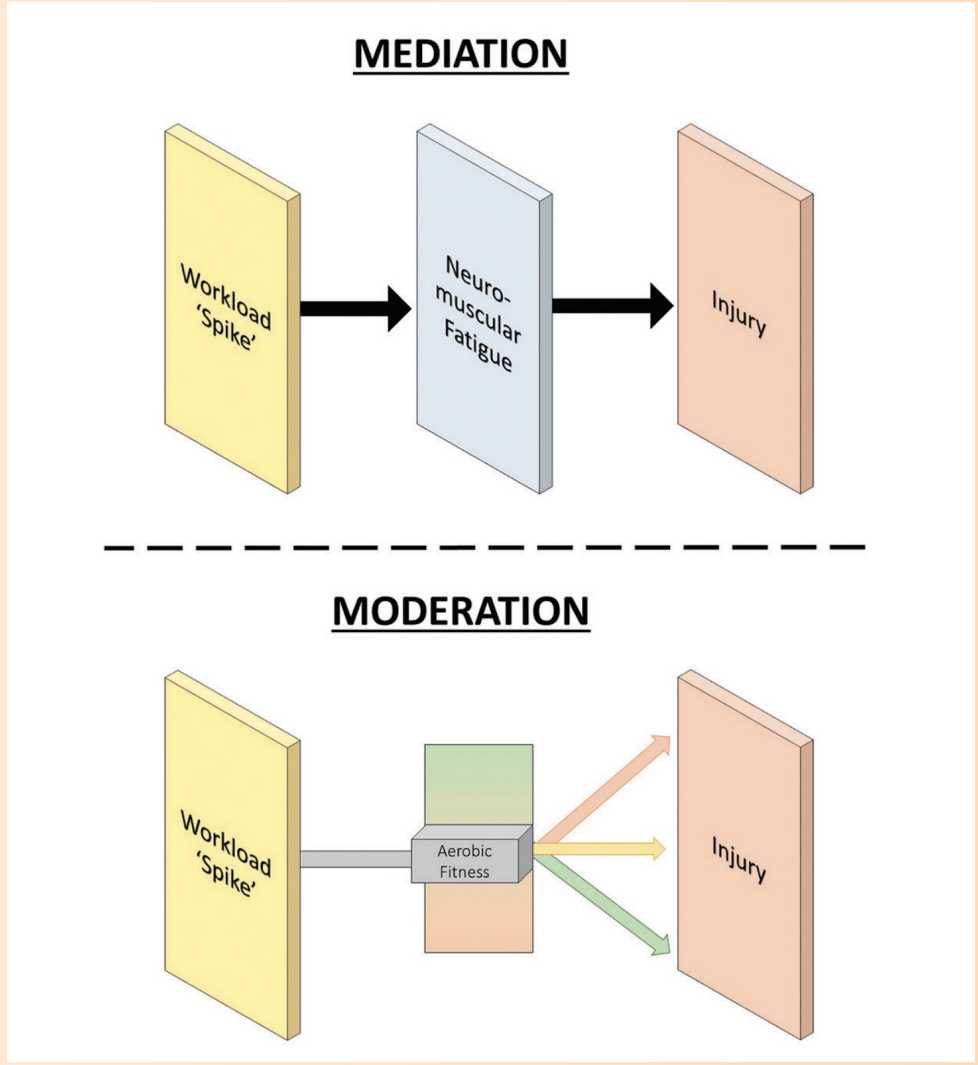
A moderator is a variable that affects the strength of the relation between the predictor and criterion variable. Moderators specify when or to whom a relation will hold. Moderating variable are typically an interaction term in statistical models.

e.g. relationship between stress and depression may differ in strength at different levels of social support. In other words, stress may be more strongly associated with depression under conditions of low social support compared to conditions of high social support.

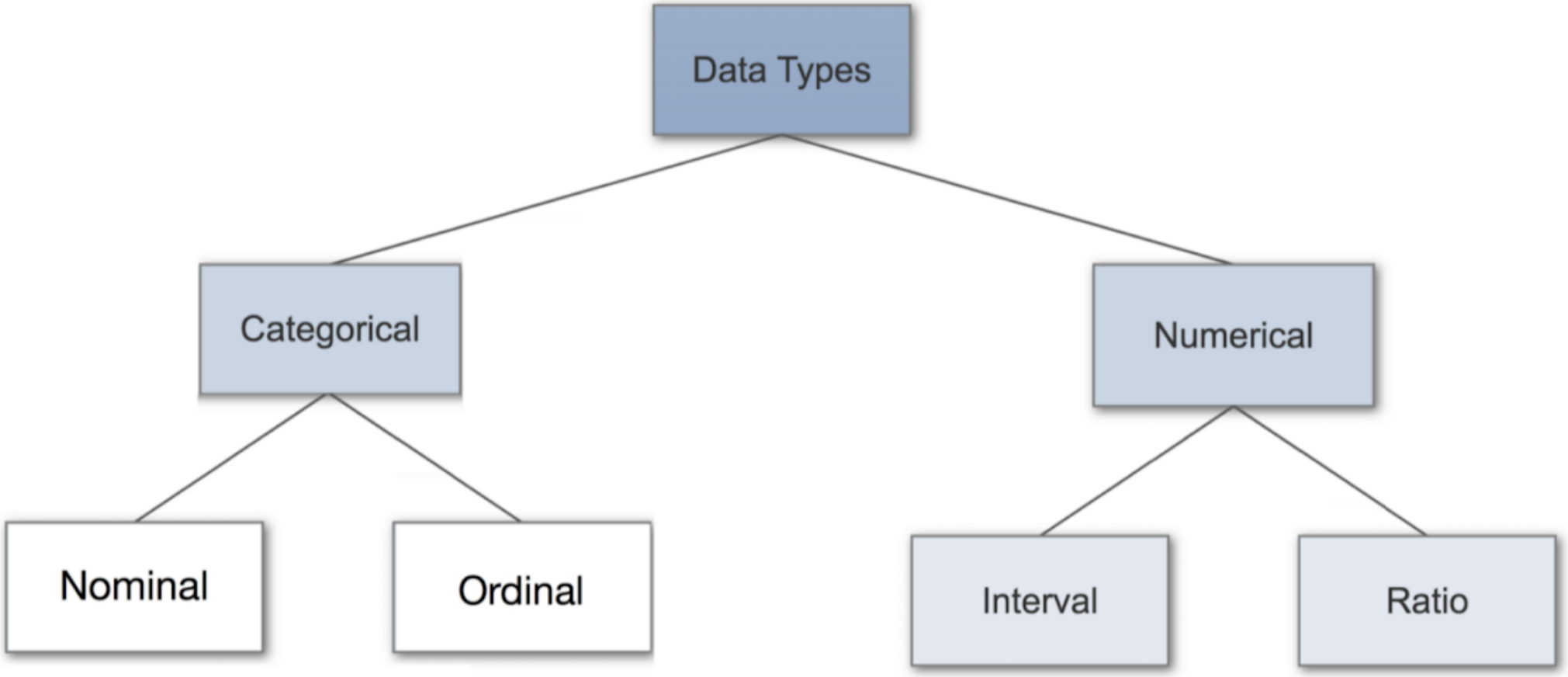


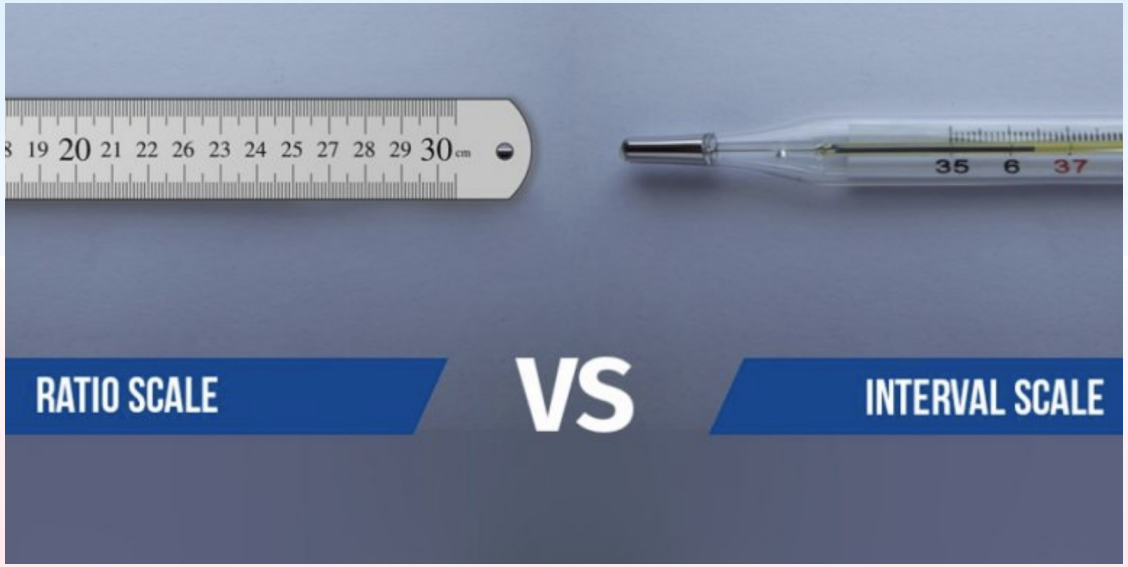
Mediation vs moderation

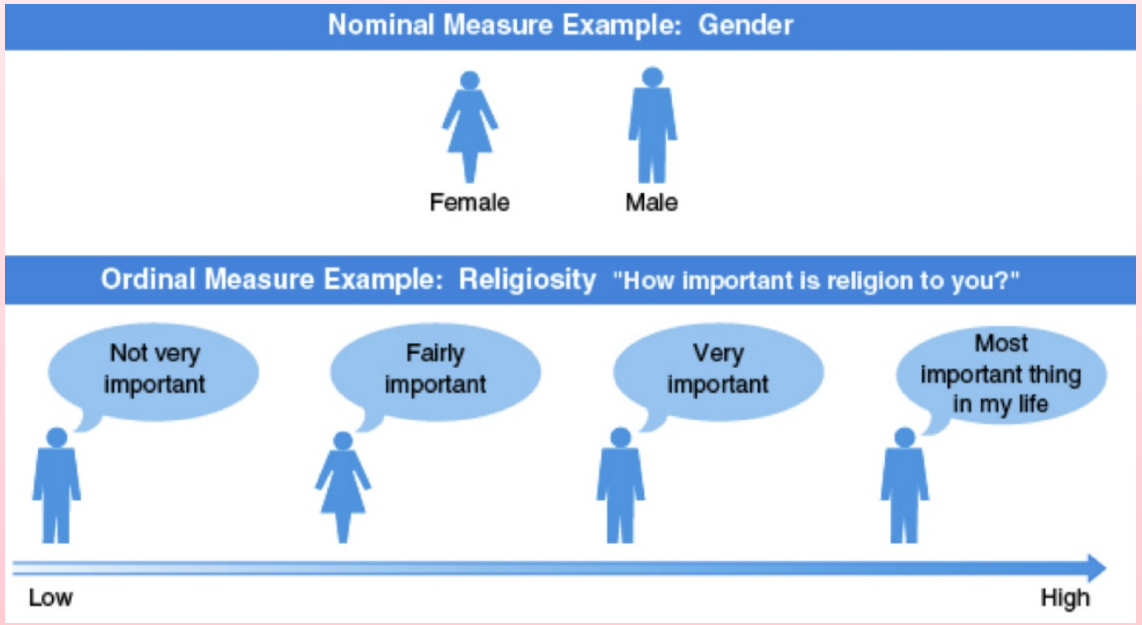
- Examples of how mediating and moderating variables explain the association between a workload spike and subsequent injury. A ME variable answers the question about why or how?
- Top, neuromuscular fatigue is shown as a mediator (ie, domino) of the effect of workload spikes on injury. In this example, a spike in workload causes an increase in neuromuscular fatigue, and an increase in neuromuscular fatigue leads to an injury. A MO variable answers the question about who or what?
- Bottom, aerobic fitness acts as a moderator (ie, dimmer switch) of the relationship between workload spikes and injury. In this case, a given spike in workload will elicit a different injury risk dependent on an individual's fitness level.

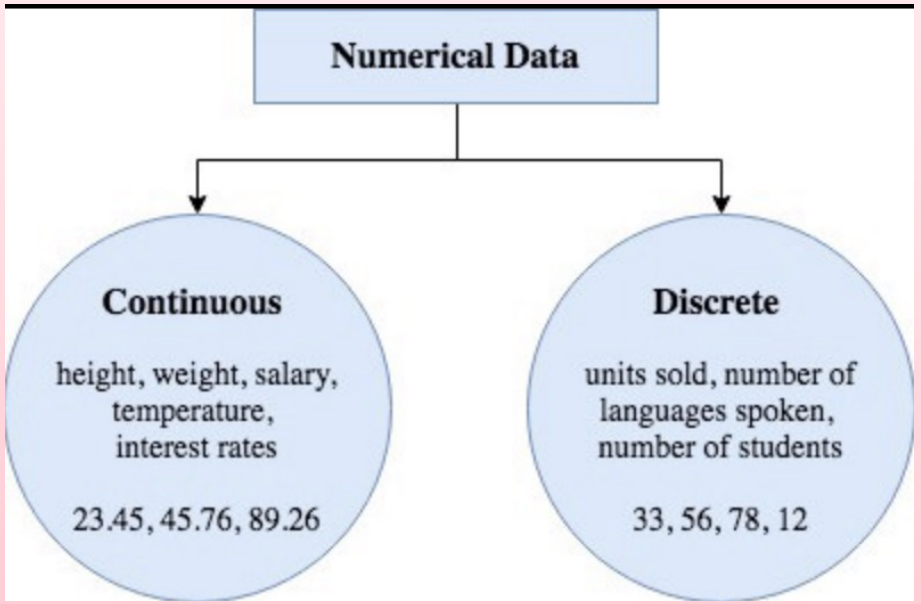


What they (variables) look like?

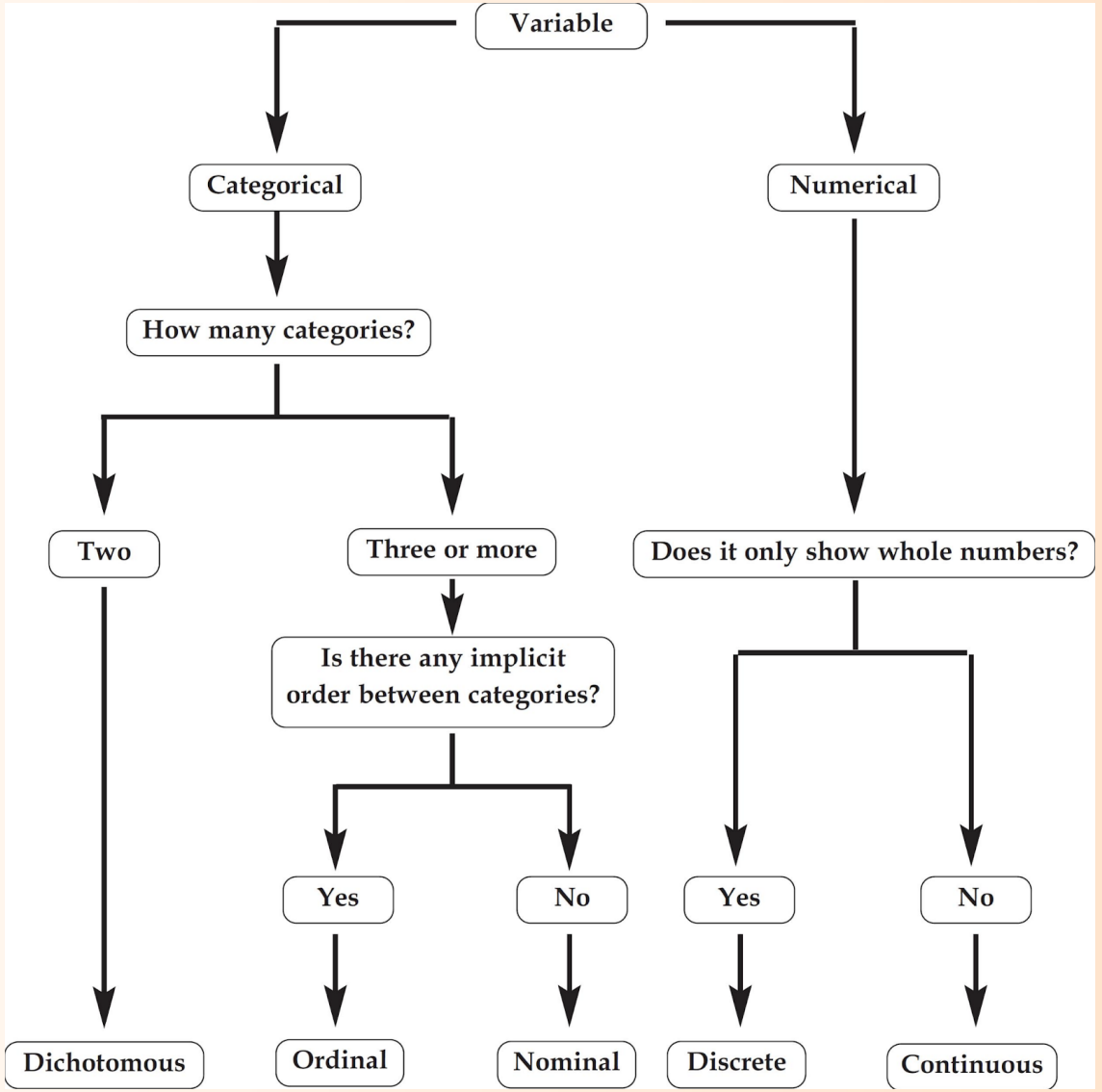



RATIO SCALE
VS
INTERVAL SCALE


Nominal Measure Example: Gender
Female
Male
Ordinal Measure Example: Religiosity "How important is religion to you?"
Not very important
Fairly important
Very important
Most important thing in my life
Low
High


Numerical Data
Continuous
height, weight, salary, temperature, interest rates
23.45, 45.76, 89.26
Discrete
units sold, number of languages spoken, number of students
33, 56, 78, 12

Classification of variables



Data Analysis: putting the currency at work

Data analysis is defined as a process of cleaning, transforming, and modeling data to discover useful information decision-making.

The purpose of Data Analysis is to extract useful information from data and making decision based upon the data analysis.

Steps in data Analysis

- Data Cleaning
  :removing duplicates/outliers.
- Data Analysis
  :statistical analysis.
- Data Interpretation
  :communicating the findings.
- Data Visualization
  :bar charts, scatter plots.

Types of Analyses (with respect to study objectives)

Descriptive

Measures of central tendency

Correlational $r_{x,y}$?

Is there a relationship between X and Y?

Measures of relationships
(correlations)

Causal $\Delta X \rightarrow \Delta Y$?

- Does a change in X cause a change in Y?

Correlation vs causation

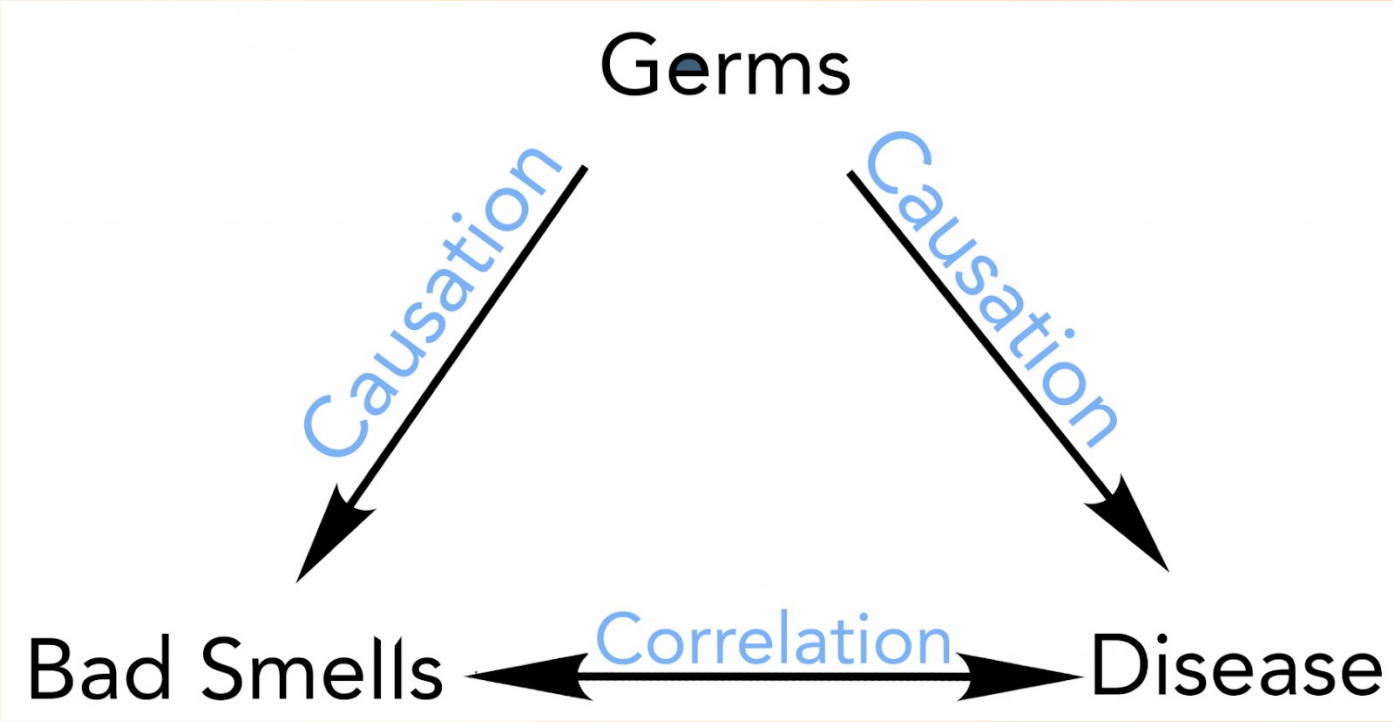
- While causation and correlation can exist at the same time, correlation does NOT imply causation. Causation explicitly applies to cases where action A causes outcome B. On the other hand, correlation is simply a relationship. Action A relates to Action B—but one event doesn't necessarily cause the other event to happen.

- Correlation and causation are often confused because the human mind likes to find patterns even when they do not exist. We often fabricate these patterns when two variables appear to be so closely associated that one is dependent on the other. That would imply a cause and effect relationship where the dependent event is the result of an independent event.

Correlation vs causation

Causation explicitly applies to cases where action A causes outcome B.

On the other hand, correlation is simply a relationship.



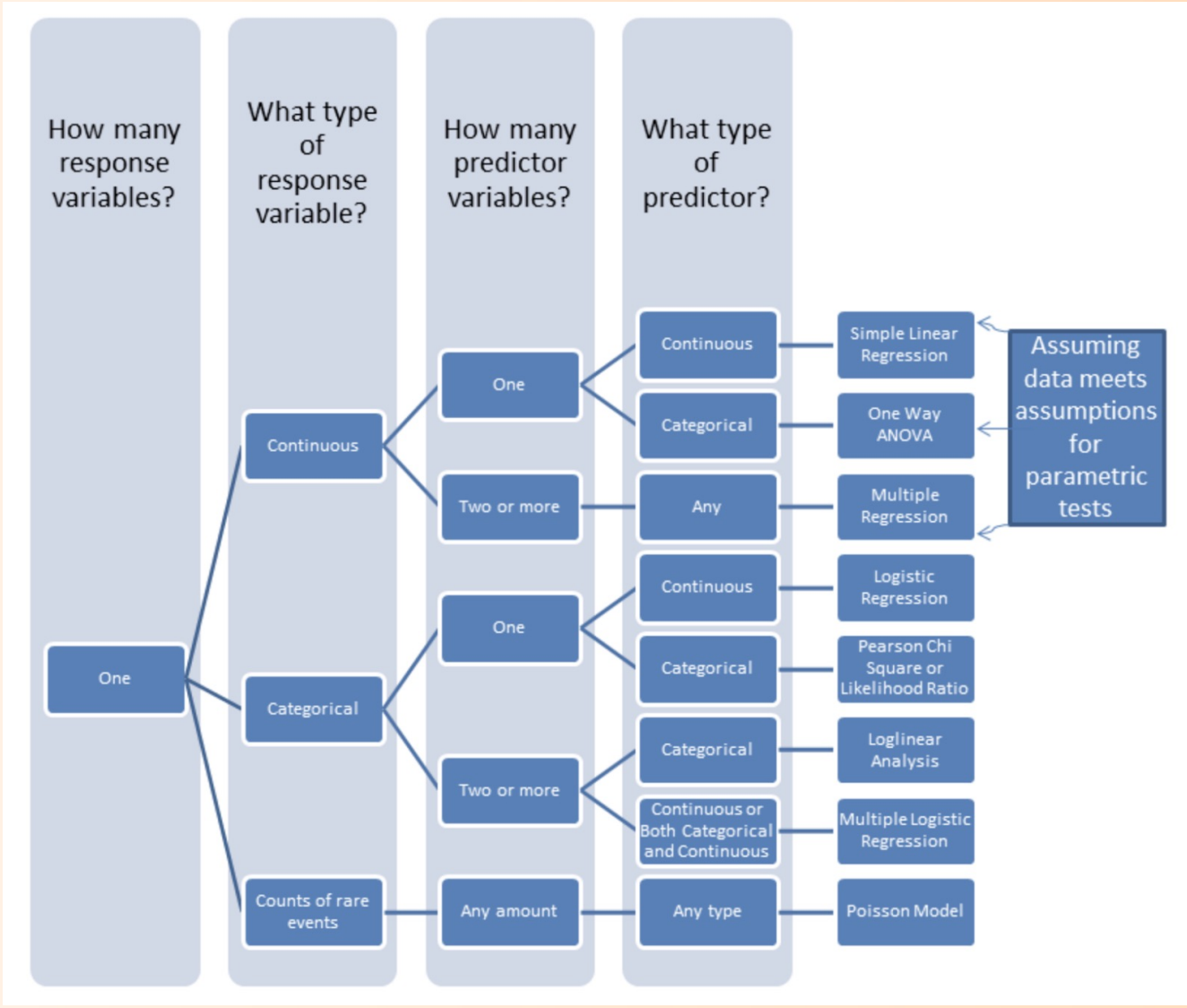
Types of Analyses

with respect to number of variables + study objectives)

Univariate Data	Bivariate Data
• involving a **single variable**	• involving **two variables**
• does not deal with causes or relationships	• deals with causes or relationships
• the major purpose of univariate analysis is to describe	• the major purpose of bivariate analysis is to explain
• central tendency - mean, mode, median • dispersion - range, variance, max, min, quartiles, standard deviation. • frequency distributions • bar graph, histogram, pie chart, line graph, box-and-whisker plot	• analysis of two variables simultaneously • correlations • comparisons, relationships, causes, explanations • tables where one variable is contingent on the values of the other variable. • independent and dependent variables
Sample question: How many of the students in the freshman class are female?	**Sample question:** Is there a relationship between the number of females in Computer Programming and their scores in Mathematics?

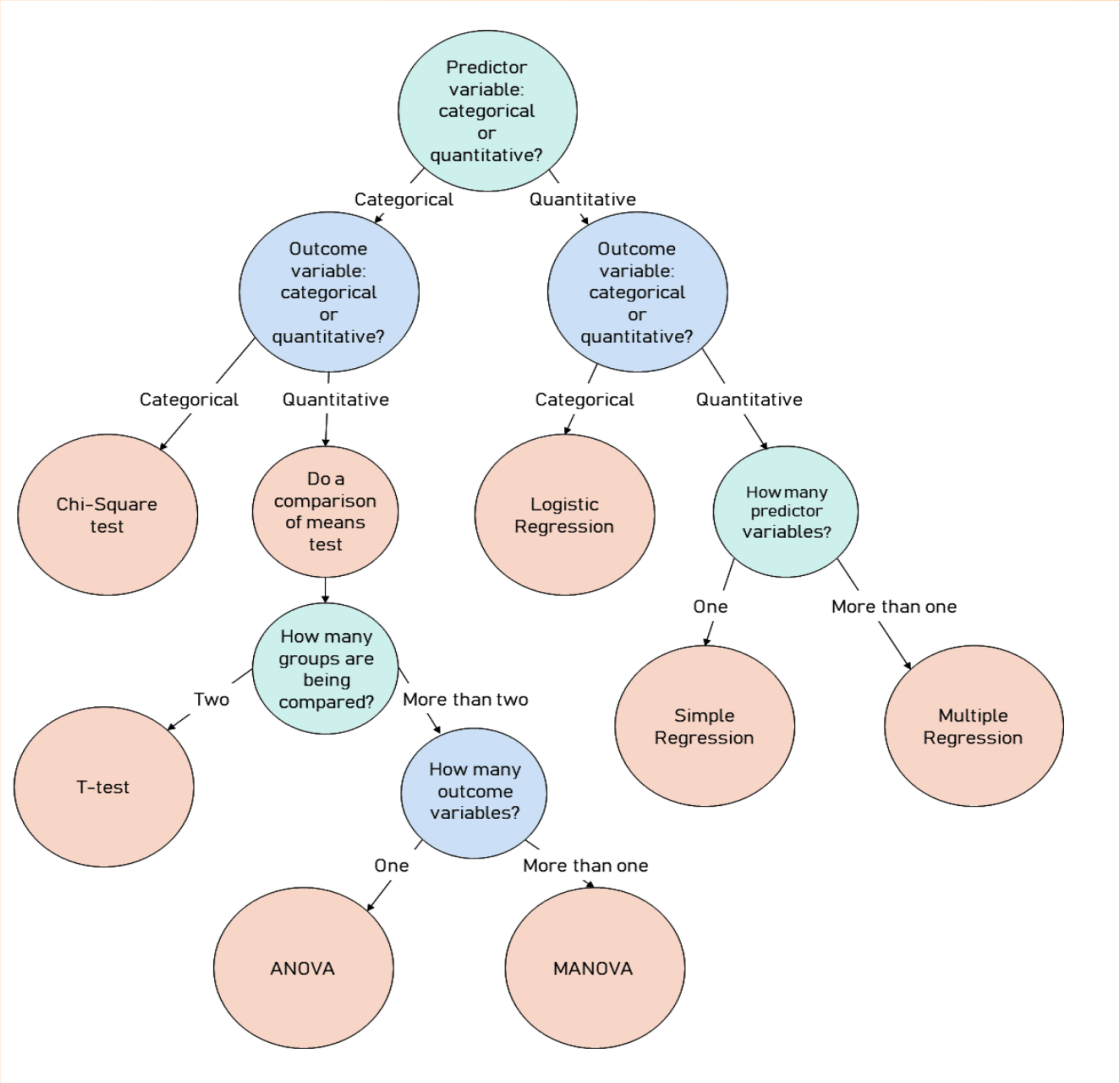
Types of regression methods (1)

- Sometimes >1 suitable method.
- Depends of the type of DV and IV.



Types of regression methods (2)

- Sometimes >1 suitable method.
- Depends of the type of DV and IV.



Multiple Regression Analysis (multivariate)

FORMULA 16.3 $$Y = a + b_1X_1 + b_2X_2$$

where b_1 = the partial slope of the linear relationship between the first independent variable and Y

b_2 = the partial slope of the linear relationship between the second independent variable and Y

Y = dependent variable (health status)

$X_{1\text{-}2}$ = predictors/ independent variables

Health Status = Age + Smoking

Types of papers

- Original research (~3500 words(excluding abstract, ref, appendix), 5 tables/figures, 30 references)
- Protocol (Typically 3500 words, 5 tables/figures, 30 references)
- Review: mini (~4000), systematic, meta-analysis:(~7000 words, 80 references)
- Commentary (~1000 words, 10 references)
- Short reports (~1000 words, 10 references)
- Letter to editor (~600 words, 5 references)

Standard structure of a manuscript

- 'AIMRDC' structure
- Title: Should reflect the:
- Key words + target population + setting + type of research
- HIV and stigmatization among university students in Accra, Ghana: a cross-sectional study

Abstract

Background: What is the problem?

Objectives: What was done?

Methods: Study type (Qualitative/cross-sectional), timeline (from June to July 2020), sample population (adolescents), number (n=), key variables, statistical approaches.

Results: key findings

Conclusion: What do the findings refer to? What could be done now?

Introduction

General description (~150 words): HIV is a global health issue…

Contextualization (~200 words): Treatment seeking for HIV can be influenced by stigma, **especially** among elderly/women)

Mechanisms (~250 words): how stigmatization affects treatment-seeking)

Rationale (~100 words): The present study addressed the research gap by…

Potential outcomes (~50 words)

Methods

Study design: sampling, data collection.

Definition of variables (Clearly identify outcome & explanatory variables, define with references)

Statistical approaches: How missing values were treated, what are the main analyses, sensitivity tests, model fitness, level of significance.

Ethical clearance: IRB no/not applicable.

Results

Sample profile

Main course of analyses

Sensitivity analyses

Model fitness

Discussion & conclusion

Main findings - ~200 words

Contrasting interpretation - ~100

Policy implication ~ 200

Strengths and limitations

Life cycle of research publication

Conceptualization (e.g. identifying the problem/hypothesis)

Study design (e.g. data collection approach)

Carrying out the research (e.g. data analysis, interpretation)

Preparing the manuscript (rationale, style, structure)

Finding the right journal/s (areas of focus)

Manuscript submission

Peer-review

Thanks